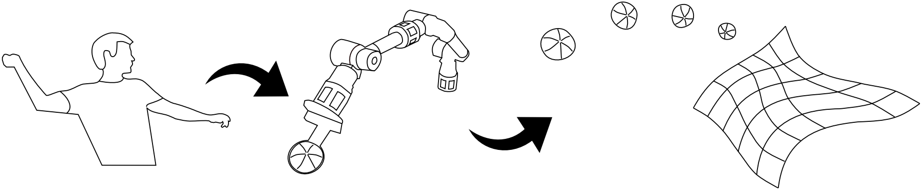
Kolloquium

Learning and Generalizing Behaviors for Robots from Human Demonstration

Alexander Fabisch (DFKI, Universität Bremen)



Gutachter: Prof. Dr. Dr. h. c. Frank Kirchner
Universität Bremen

Prof. Constantin A. Rothkopf, PhD
Technische Universität Darmstadt

3. Dezember 2020

1. Motivation
2. State of the Art
3. Contributions
   - Conceptual Framework for Automatic Robot Behavior Learning
   - Imitation with Automatic Embodiment Mapping
   - Sample-Efficient Contextual Policy Search
4. Discussion
5. Outlook

Corresponding publications

Motivation

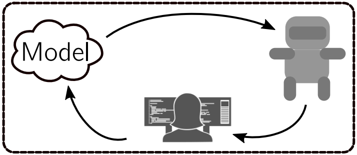
Why Should a Robot Learn?

Example:

Why Should a Robot Learn?

Example:

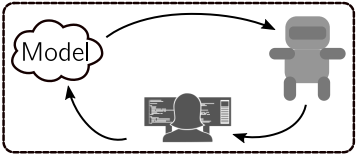
- Many different tasks
- We can learn to solve a wide range of tasks
- Learning replaces the feedback loop:



Why Should a Robot Learn?

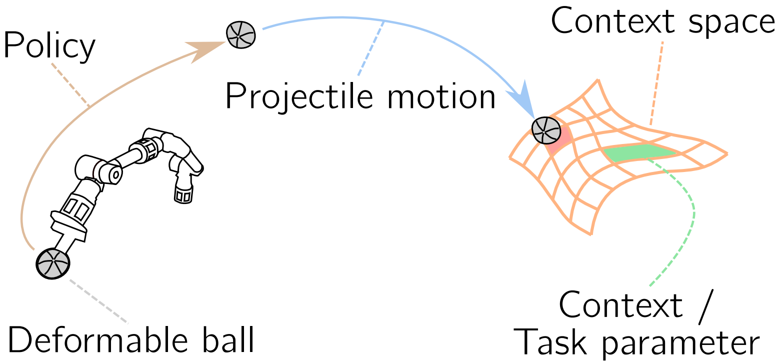
Example:

- Many different tasks
- We can learn to solve a wide range of tasks
- Learning replaces the feedback loop:



Learning should be a common tool for behavior generation.

Example: Ball Throwing



- Standard deviation of 4.5–7 cm per throw
  (noise, inaccurate model)
- People don't compute projectile motions
  (heuristic, easier to implement)

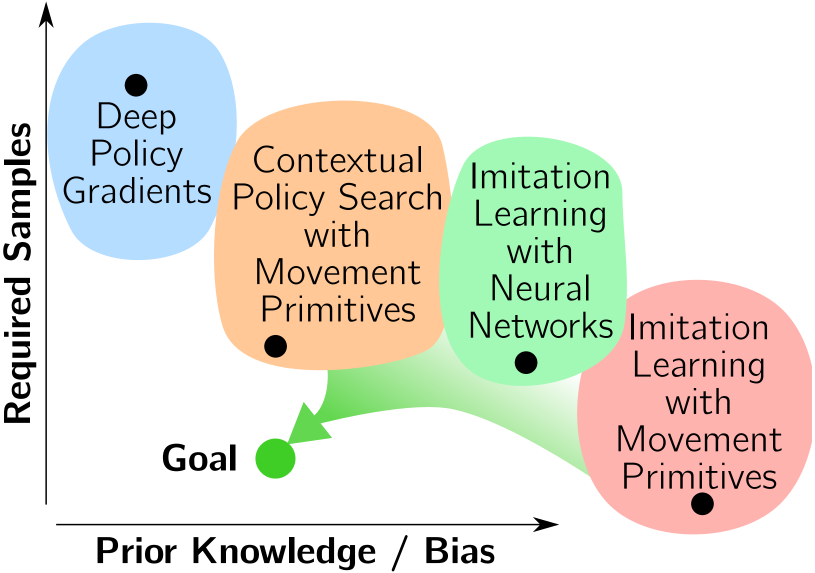
Learning on Real Robots

- Robots can break (things)
- Wear and tear
- Hardware is expensive
- Maintenance

➲ Goal: minimize interaction with the environment

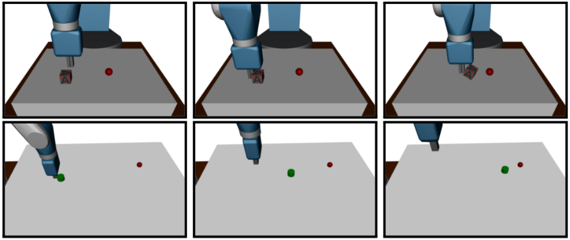
State of the Art

State of the Art: Overview



Hindsight Experience Replay (HER)

(Andrychowicz et al. 2017)



- Baseline and based on: Deep Deterministic Policy Gradients (DDPG) (Lillicrap et al. 2016)
- Policy: neural network
- Return: binary, goal tolerance 7 / 20 cm
- 160,000 episodes, > 90% success rate

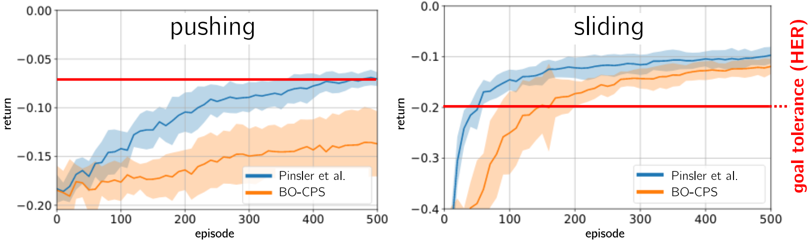
Hindsight Experience Replay (HER)

(Andrychowicz et al. 2017)

- Baseline and based on: Deep Deterministic Policy Gradients (DDPG) (Lillicrap et al. 2016)
- Policy: neural network
- Return: binary, goal tolerance 7 / 20 cm
- 160,000 episodes, $> 90\%$ success rate

Contextual Policy Search (CPS)

(Pinsler et al. 2019)



- Baseline and based on: Bayesian Optimization for CPS (BO-CPS) (Metzen et al. 2015)
- Policy: Dynamical Movement Primitives
- Return: negative distance to target
- Orders of magnitude faster than HER

Contextual Policy Search (CPS)

Upper-level policy

$$\pi_{\boldsymbol{\omega}}(\boldsymbol{\theta}|\boldsymbol{s})$$

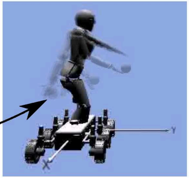
Context

$$\boldsymbol{s} = \begin{pmatrix} s_1 \\ s_2 \\ \vdots \\ s_{n_s} \end{pmatrix}$$

Parameters

$$\pi_{\boldsymbol{\theta}}(\boldsymbol{x}_{t-1}, t) = \boldsymbol{x}_t$$

Control policy



Dynamical Movement Primitive (DMP)

DMPs (Mülling et al. 2013) define trajectories via

$$\boldsymbol{x}_{t'} = \pi_{\boldsymbol{w},\boldsymbol{v}}(\boldsymbol{x}_t, t)$$

- Weights $\boldsymbol{w}$ define the shape
- Metaparameters $\boldsymbol{v} = (\boldsymbol{x}_0, \boldsymbol{g}, \dot{\boldsymbol{g}}, \tau)$:
  - $\boldsymbol{x}_0$: initial position
  - $\boldsymbol{g}, \dot{\boldsymbol{g}}$: final position, velocity
  - τ: duration
- *Stable* trajectory generators
- Can be used for imitation learning (IL)

Summary

- Deep RL is not sample-efficient **yet**
- Contextual Policy Search (CPS)
  - robust
  - sample-efficient
  - needs good initialization
- Dynamical Movement Primitive (DMP)
  - initialized from imitation learning
  - domain-specific policy representation

Contributions

Challenges / Goals / Contributions

Challenges

01

Behavior learning is not a common tool	Sample efficiency	Learned behaviors do not generalize	Variety of considered problems is limited

Goals

02

Reduce required expert knowledge	Sample efficiency (100-300 episodes)	Generalization over context space	Evaluation on various tasks and systems

03

Contributions

04

- Conceptual framework for automatic robot behavior learning
  - Automatic embodiment mapping for imitation
  - Sample-efficient contextual policy search

Software: BOLeRo

Corresponding publications

Challenges / Goals / Contributions

Challenges

Behavior learning is not a common tool	Sample efficiency	Learned behaviors do not generalize	Variety of considered problems is limited

Goals

Reduce required expert knowledge	Sample efficiency (100-300 episodes)	Generalization over context space	Evaluation on various tasks and systems

Contributions

- Conceptual framework for automatic robot behavior learning
- Automatic embodiment mapping for imitation
- Sample-efficient contextual policy search

Software: BOLeRo

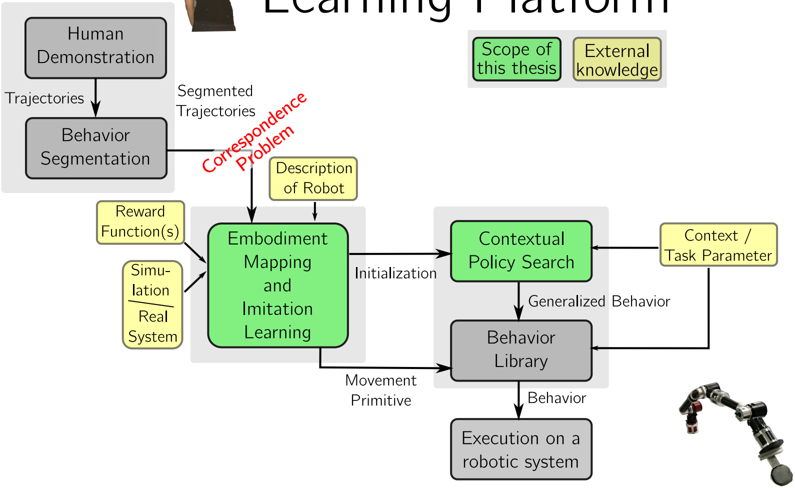
01 02 03 04

Corresponding publications

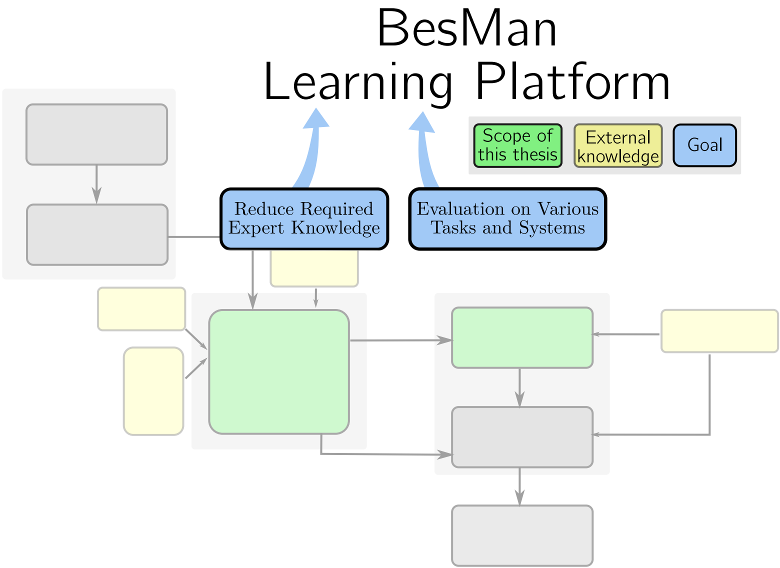
Contribution: Conceptual Framework for Automatic Robot Behavior Learning

Framework

BesMan Learning Platform



Framework


BesMan
Learning Platform
Scope of this thesis
External knowledge
Goal
Reduce Required Expert Knowledge
Evaluation on Various Tasks and Systems

Contributions

Challenges

Behavior learning is not a common tool

Sample efficiency

Learned behaviors do not generalize

Variety of considered problems is limited

Goals

Reduce required expert knowledge

Sample efficiency (100-300 episodes)

Generalization over context space

Evaluation on various tasks and systems

Contributions

Conceptual framework for automatic robot behavior learning

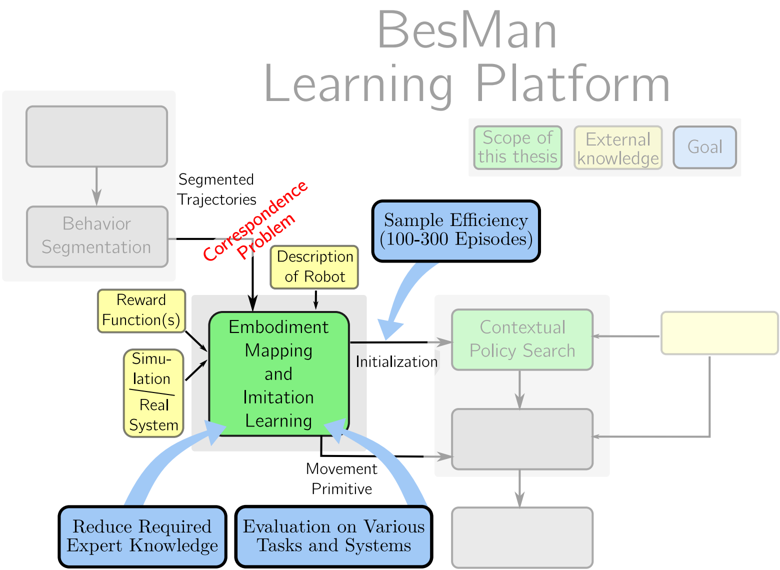
Automatic embodiment mapping for imitation

Sample-efficient contextual policy search

Software: BOLeRo

Contribution: Imitation with Automatic Embodiment Mapping

Imitation with Embodiment Mapping



Automatic Embodiment Mapping

Task-agnostic*

1. Global trajectory optimization
2. Local pose optimization
3. Spatial and temporal scaling

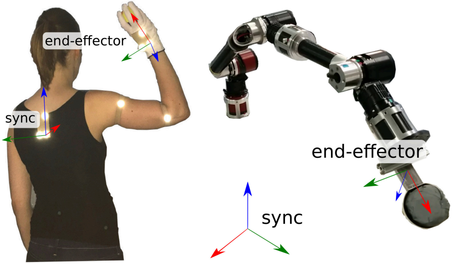
*Only considers kinematics.

Task-specific^

4. Refinement with policy search

^Adaptation to demonstration or to a similar new task

Global Trajectory Optimization

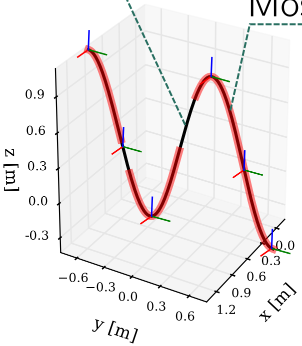


- End-effector trajectories are defined in **synchronization frames**
- These can be optimized for the target system
- Takes into account:
  reachability, joint speeds
- Does not compensate for:
  low torque, fingers vs. open scoop

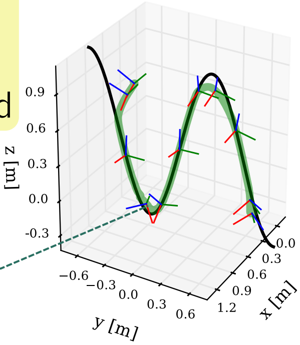
Local Pose Optimization

Example: follow sine with Kuka iiwa

Mostly not reachable (red)



Solution:
Optimize joint angles to minimize
- pose error
- joint speed



Closest reachable poses

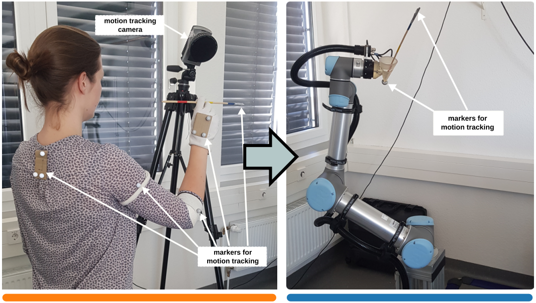
Local Pose Optimization

Evaluation of Task-Agnostic Part

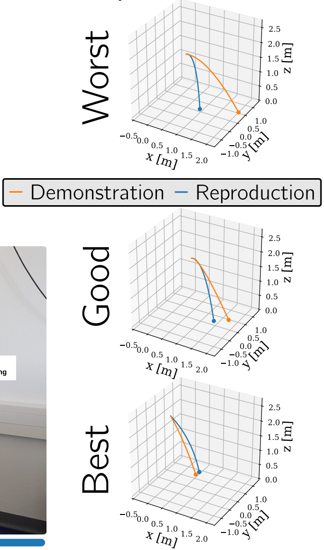
Evaluation of Task-Agnostic Part

03

- Qualitative comparison of **stick trajectories** from 33 demonstrations
- 27/33 were transferred
- Further refinement is required



Reproductions:

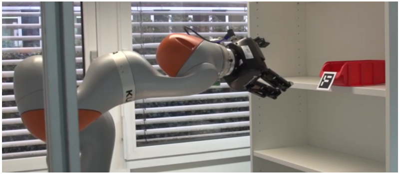


Refinement with Policy Search

- Requires task description (reward)
- Existing methods: REPS, CMA-ES, Bayesian optimization

Episodes (in simulation):

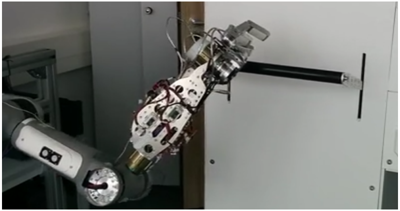
Refinement with Policy Search



- Requires task description (reward)
- Existing methods: REPS, CMA-ES, Bayesian optimization

Episodes (in simulation): 50–100 (grasp)

Refinement with Policy Search



- Requires task description (reward)
- Existing methods: REPS, CMA-ES, Bayesian optimization

Episodes (in simulation): 50–100 (grasp), ca. 300 (pull)

Refinement with Policy Search

- Requires task description (reward)
- Existing methods: REPS, CMA-ES, Bayesian optimization

Episodes (in simulation): 50–100 (grasp), ca. 300 (pull), 50 (throw)

Summary (Embodiment Mapping)

- Contribution: procedure for automatic embodiment mapping
- Output: robot-specific movement primitive
- Task-specific refinement is required

02 The BesMan Learning Platform for Automated Robot Skill Learning

L. Gutzeit, **A. Fabisch**, M. Otto, J.H. Metzen, J. Hansen, F. Kirchner, E.A. Kirchner (main author)
Journal: Frontiers in Robotics and AI

03 Automated Robot Skill Learning from Demonstration for Various Robot Systems

L. Gutzeit, **A. Fabisch**, C. Petzoldt, H. Wiese, F. Kirchner (main author)
Conference: KI: Advances in Artificial Intelligence

04 A Comparison of Policy Search in Joint Space and Cartesian Space for Refinement of Skills

A. Fabisch (main author)
Conference: International Conference on Robotics in Alpe-Adria-Danube Region (RAAD)

Evaluation of Learning Platform

My contribution:
Embodiment mapping and IL

Average over
10 subjects x 3 datasets

Step	Time / 8 throws	Automized	Required knowledge
Motion capture	2:03 min	✗	Marker setup
Marker labeling	4:58 min	✓	
Behavior segmentation	0:44 min	✓	
Imitation learning	4:20 min	✓	Robot model
Transferability approach	85 min	(✓)	Reward, simulation

- Mostly automated workflow
- Interaction with real world is costly (50 episodes)

Summary (Framework)

- Contributions:
  1. embodiment mapping and CPS modules
  2. integration of components
  3. evaluation
- Mostly automated workflow
- Some knowledge is still required (markers, robot model, reward)

The BesMan Learning Platform for Automated Robot Skill Learning
L. Gutzeit, **A. Fabisch**, M. Otto, J.H. Metzen, J. Hansen, F. Kirchner, E.A. Kirchner (main author)
Journal: Frontiers in Robotics and AI

Contributions

Challenges

- Behavior learning is not a common tool
- Sample efficiency
- Learned behaviors do not generalize
- Variety of considered problems is limited

Goals

- Reduce required expert knowledge
- Sample efficiency (100-300 episodes)
- Generalization over context space
- Evaluation on various tasks and systems

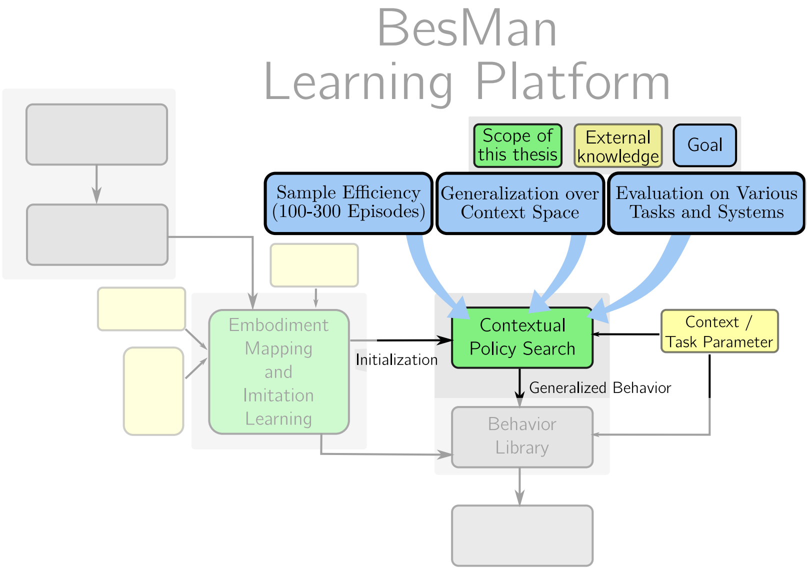
Contributions

- Conceptual framework for automatic robot behavior learning
- Automatic embodiment mapping for imitation
- Sample-efficient contextual policy search

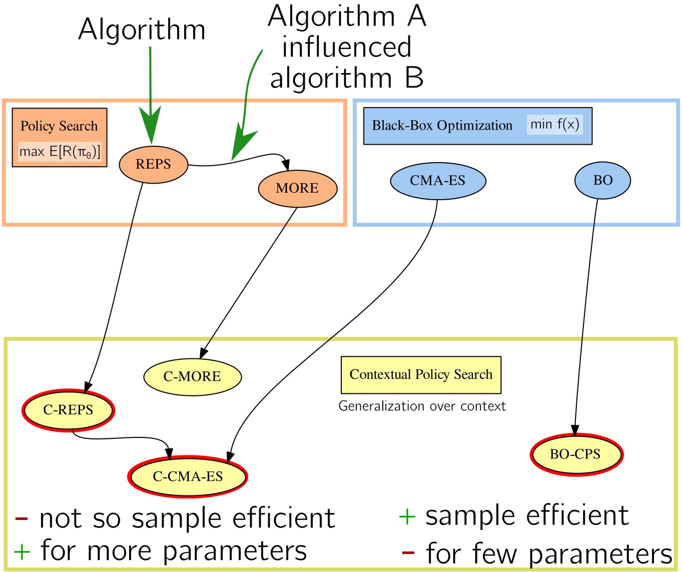
Software: BOLeRo

Contribution:
Sample-Efficient
Contextual Policy Search

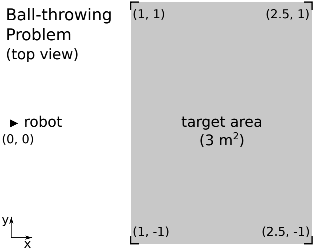
Sample-Efficient Contextual Policy Search



State of the Art (CPS)

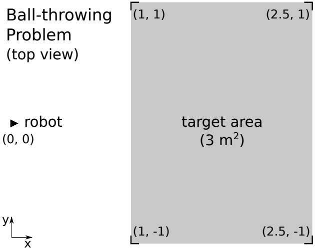


Ball Throwing



- Easy to understand
- Not too easy for RL

Ball Throwing



- Easy to understand
- Not too easy for RL

Contributions to CPS (Part 1)

Contribution

C-REPS

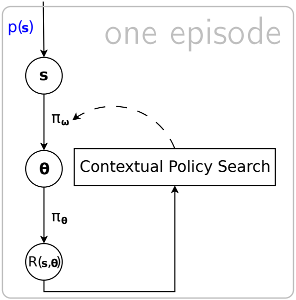
06

Active Context Selection

Sample efficiency: +33.3%
Performance: +33.5%
Episodes: > 8000
Impact beyond CPS

Active Context Selection

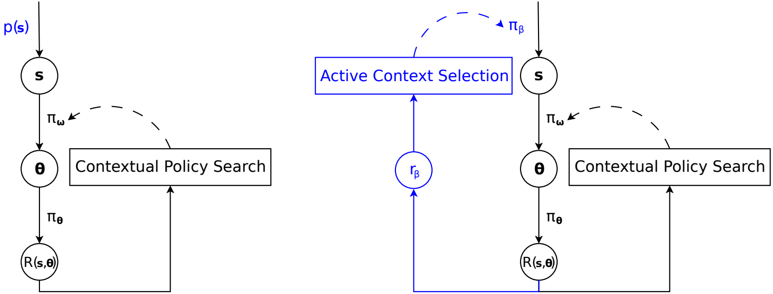
Random context exploration during training



- $p(\boldsymbol{s})$: context distribution

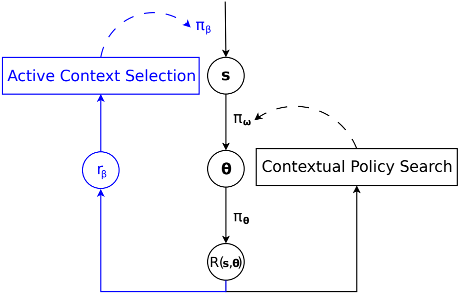
Active Context Selection

Goal: increase sample efficiency by context selection



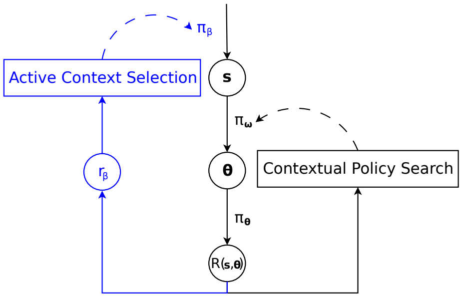
- $p(\mathbf{s})$: context distribution
- π_β: policy to select context

Active Context Selection



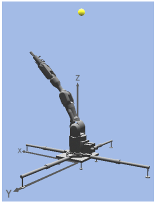
- Goal: select $\boldsymbol{s}_t \in \mathcal{S}$ to maximize $\mathbb{E}\left[R_t - R_{t-1}\right]$
- Idea: model context selection as *non-stationary multi-armed bandit problem*
- Algorithm: Discounted Upper Confidence Bound (D-UCB)

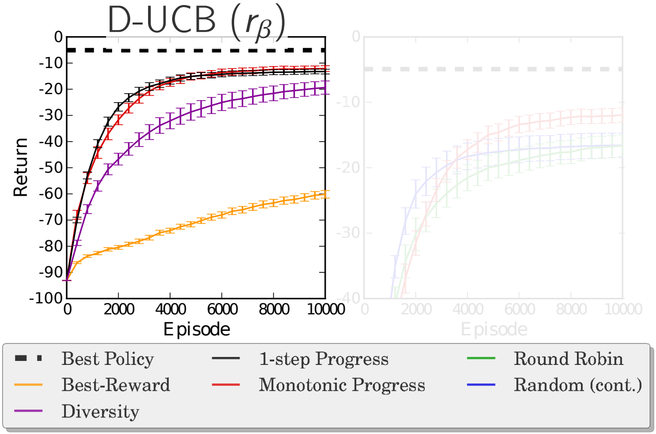
Active Context Selection



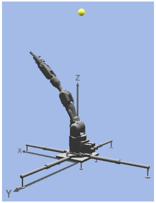
- Reward learning progress (r_β)
- Candidate heuristics for r_β:
  - 1-step progress: difference of successive returns
  - Monotonic progress: maximum of 0 and 1-step progress
  - Best reward: use return directly
  - Diversity: negative return

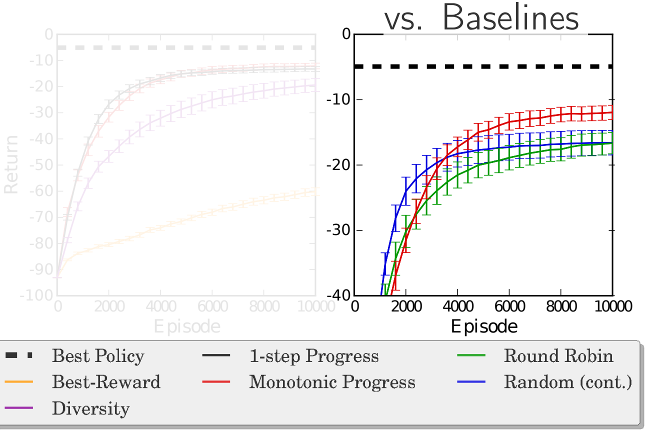
Active Context Selection for C-REPS



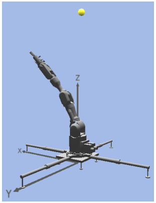


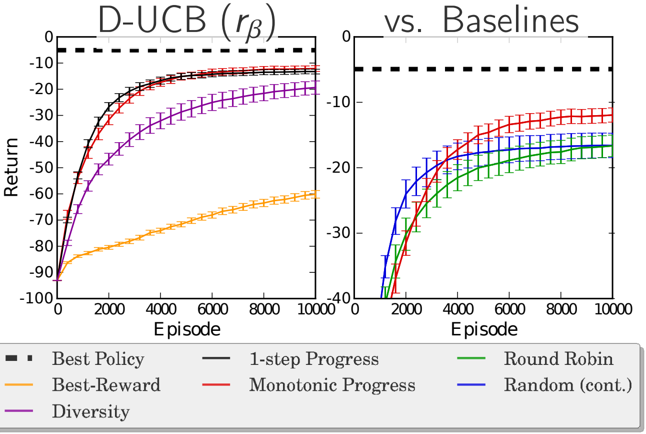
Active Context Selection for C-REPS



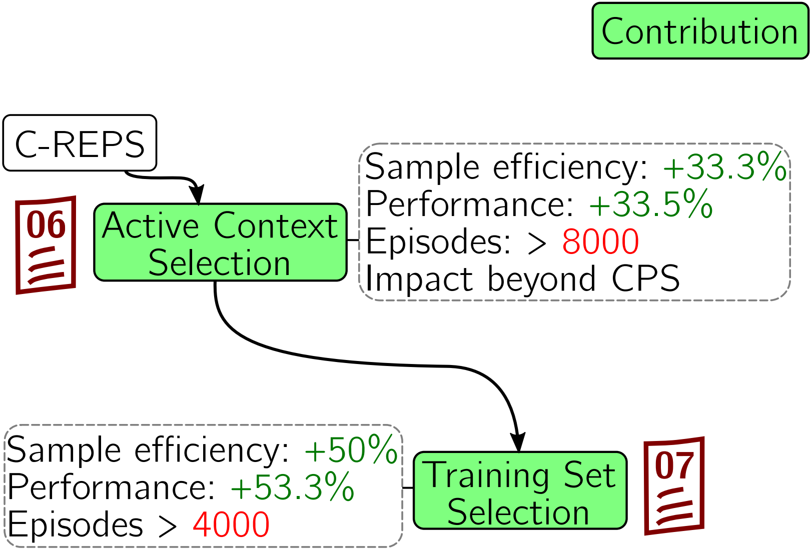


Active Context Selection for C-REPS

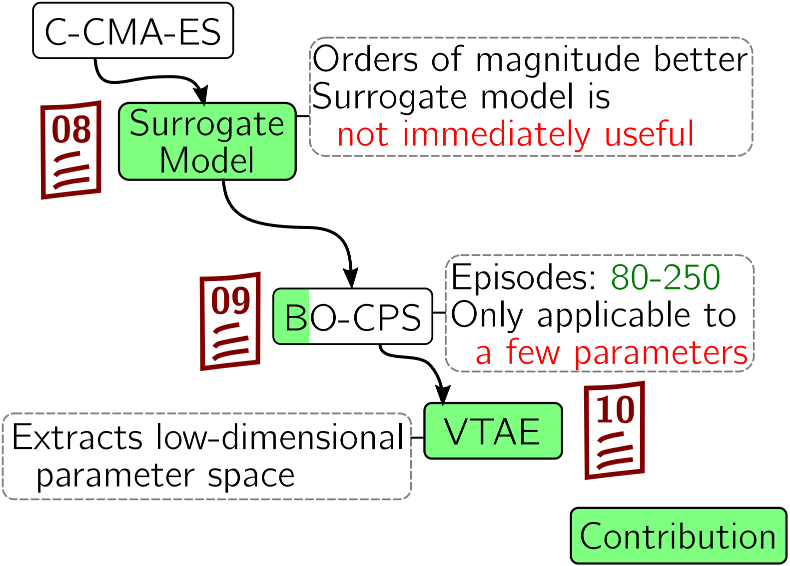




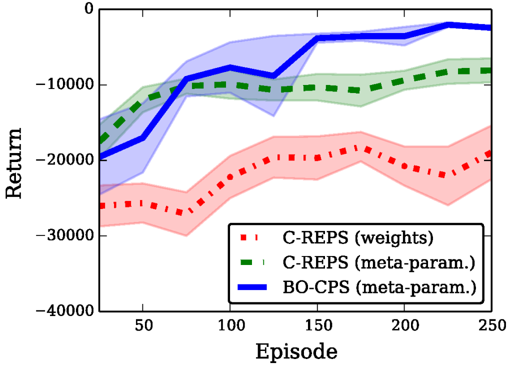
Contributions to CPS (Part 2)



Contributions to CPS (Part 3)

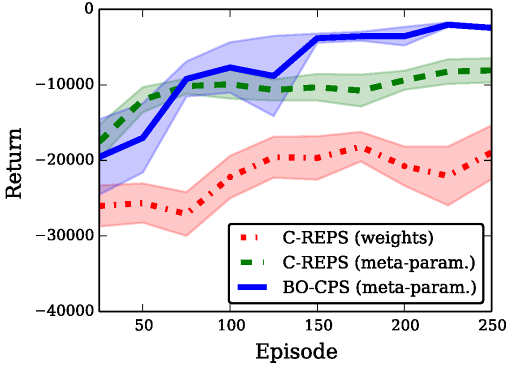


BO-CPS—Results (Simulation)



Metaparameters, manually selected: $(\boldsymbol{g}_0, \tau)$

BO-CPS—Results (Simulation)



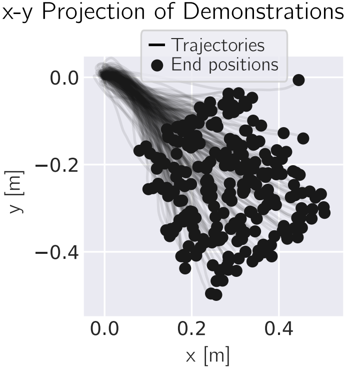
Metaparameters, manually selected: $(\boldsymbol{g}_0, \tau)$

J. Hansen (2015) obtained similar results in 80 episodes on the real robot (Master's thesis)

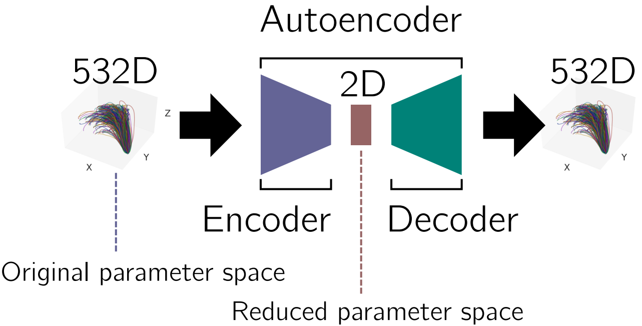
Dataset: Grasping Motions

Research Question

Can we find a small set of parameters automatically?

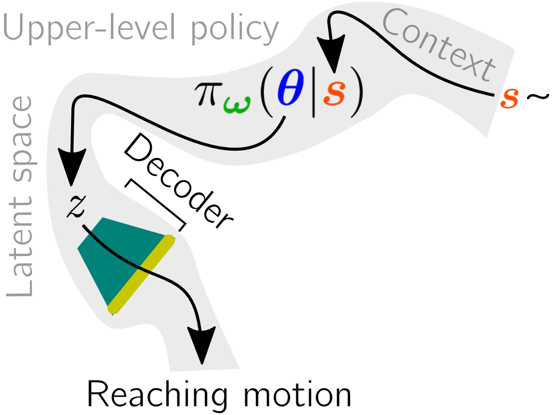


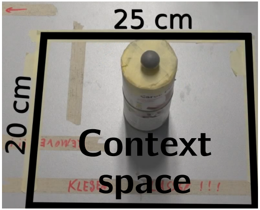
VTAE: Variational Trajectory Autoencoder



- Represents only samples that are close to training set
- VTAE: "Trajectory layer" encourages smoothness

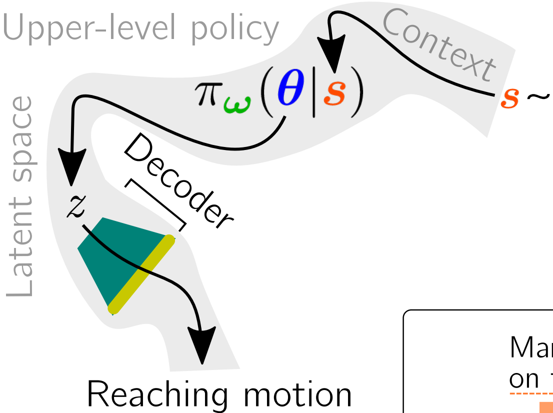
Experiment: Grasping

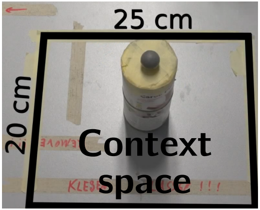


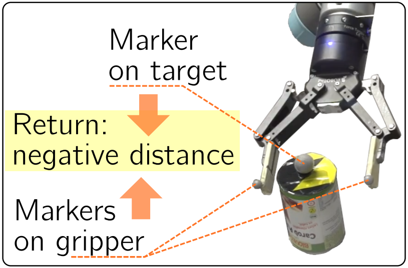


- Budget: 250 episodes

Experiment: Grasping

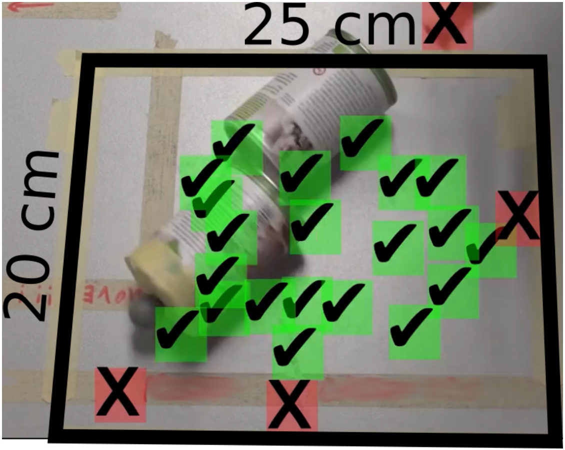




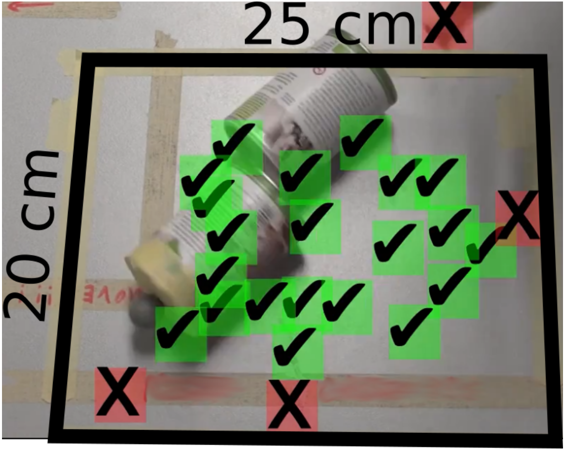


- Budget: 250 episodes

Final Policy with BO-CPS



Final Policy with BO-CPS



Summary

- Most success: reward model + manifold learning
- We can apply BO-CPS to throwing and grasping
- Active context selection led to follow-up works

Contributions:

06 Active Contextual Policy Search

A. Fabisch, J.H. Metzen (main author)
Journal: Journal of Machine Learning Research

07 Accounting for Task-Difficulty in Active Multi-Task Robot Control Learning

A. Fabisch, J.H. Metzen, M.M. Krell, F. Kirchner (main author)
Journal: Künstliche Intelligenz

Summary

- Most success: reward model + manifold learning
- We can apply BO-CPS to throwing and grasping
- Active context selection led to follow-up works

Contributions:

08 Empirical Evaluation of Contextual Policy Search with a Comparison-based Surrogate Model and Active Covariance Matrix Adaptation

A. Fabisch (main author)
Conference: Genetic and Evolutionary Computation Conference (GECCO)

09 Bayesian Optimization for Contextual Policy Search

J.H. Metzen, **A. Fabisch**, J. Hansen (co-author)
Workshop: Machine Learning in Planning and Control of Robot Motion (at IROS)

10 Variational Trajectory Autoencoder for Sample-Efficient Policy Search

A. Fabisch, F. Kirchner (main author)
Conference: Conference on Robot Learning (submitted)

Discussion

Contributions

Challenges

- Behavior learning is not a common tool
- Sample efficiency
- Learned behaviors do not generalize
- Variety of considered problems is limited

Goals

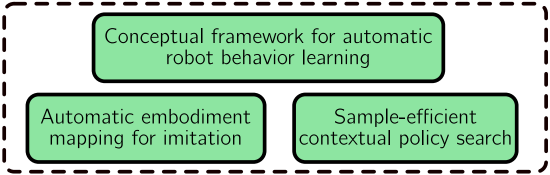
- Reduce required expert knowledge
- Sample efficiency (100-300 episodes)
- Generalization over context space
- Evaluation on various tasks and systems

Contributions

- Conceptual framework for automatic robot behavior learning
- Automatic embodiment mapping for imitation
- Sample-efficient contextual policy search

Software: BOLeRo

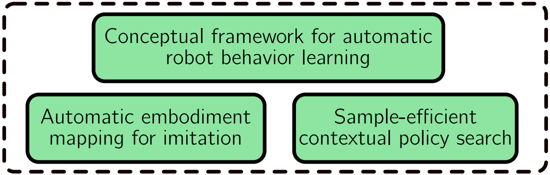
More Contributions



Software: BOLeRo

- Review of behavior learning for robots (problems, algorithms, alternatives)

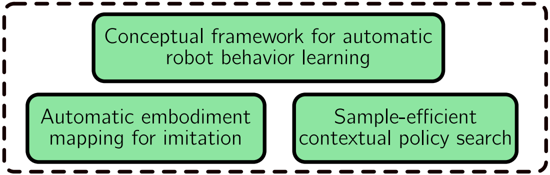
More Contributions



Software: BOLeRo

- Review of behavior learning for robots
- Evaluation of embodiment mapping on 697 demonstrations, learning platform on 240 demonstrations
- Experiments on 4 real and 7 simulated robots

More Contributions



Software: BOLeRo

- Review of behavior learning for robots
- Evaluation of embodiment mapping on 697 demonstrations, learning platform on 240 demonstrations
- Experiments on 4 real and 7 simulated robots
- Models
  - ➲ VTAE
  - ➲ PUBSVE (estimates upper boundary of data)

Outlook

Next Steps

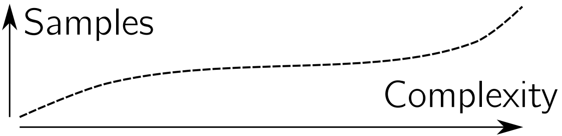
- Combination of trajectory generators with computer vision
- Exploration of methods for reward generation

Next Steps

- Combination of trajectory generators with computer vision
- Exploration of methods for reward generation

Questions

- What should we learn?
- Which inductive biases should we use?
- How should we measure progress?



Appendix

Publications I

Introduction

01 A Survey of Behavior Learning Applications in Robotics—State of the Art and Perspectives

A. Fabisch, C. Petzoldt, M. Otto, F. Kirchner (main author)
Journal: International Journal of Robotics Research (submitted)

Automatic Embodiment Mapping

02 The BesMan Learning Platform for Automated Robot Skill Learning

L. Gutzeit, **A. Fabisch**, M. Otto, J.H. Metzen, J. Hansen, F. Kirchner, E.A. Kirchner (main author)
Journal: Frontiers in Robotics and AI

03 Automated Robot Skill Learning from Demonstration for Various Robot Systems

L. Gutzeit, **A. Fabisch**, C. Petzoldt, H. Wiese, F. Kirchner (main author)
Conference: KI: Advances in Artificial Intelligence

04 A Comparison of Policy Search in Joint Space and Cartesian Space for Refinement of Skills

A. Fabisch (main author)
Conference: International Conference on Robotics in Alpe-Adria-Danube Region (RAAD)

Publications II

Contextual Policy Search

05 Learning in Compressed Space

A. Fabisch, Y. Kassahun, H. Wöhrle, and F. Kirchner (main author)
Journal: Neural Networks

06 Active Contextual Policy Search

A. Fabisch, J.H. Metzen (main author)
Journal: Journal of Machine Learning Research

07 Accounting for Task-Difficulty in Active Multi-Task Robot Control Learning

A. Fabisch, J.H. Metzen, M.M. Krell, F. Kirchner (main author)
Journal: Künstliche Intelligenz

08 Empirical Evaluation of Contextual Policy Search with a Comparison-based Surrogate Model and Active Covariance Matrix Adaptation

A. Fabisch (main author)
Conference: Genetic and Evolutionary Computation Conference (GECCO)

09 Bayesian Optimization for Contextual Policy Search

J.H. Metzen, **A. Fabisch**, J. Hansen (co-author)
Workshop: Machine Learning in Planning and Control of Robot Motion (at IROS)

10 Variational Trajectory Autoencoder for Sample-Efficient Policy Search

A. Fabisch, F. Kirchner (main author)
Conference: Conference on Robot Learning (submitted)

Publications III

Framework

11 Towards Learning of Generic Skills for Robotic Manipulation

J.H. Metzen, **A. Fabisch**, L. Senger, J. de Gea Fernandez, E.A. Kirchner (co-author)
Journal: Künstliche Intelligenz

12 BOLeRo: Behavior Optimization and Learning for Robots

A. Fabisch, M. Langosz, F. Kirchner (main author)
Journal: International Journal of Advanced Robotic Systems

Literature I

Abdolmaleki, Abbas, Bob Price, Nuno Lau, Luís Paulo Reis, and Gerhard Neumann (2017). "Contextual Covariance Matrix Adaptation Evolutionary Strategies". In: *International Joint Conference on Artificial Intelligence (IJCAI)*. Ed. by Carles Sierra, pp. 1378–1385. DOI: 10.24963/ijcai.2017/191.

Adi-Japha, Esther, Avi Karni, Ariel Parnes, Iris Loewenschuss, and Eli Vakil (2008). "A Shift in Task Routines During the Learning of a Motor Skill: Group-Averaged Data May Mask Critical Phases in the Individuals' Acquisition of Skilled Performance". In: *Journal of Experimental Psychology: Learning, Memory, and Cognition* 24, pp. 1544–1551.

Andrychowicz, Marcin, Filip Wolski, Alex Ray, Jonas Schneider, Rachel Fong, Peter Welinder, Bob McGrew, Josh Tobin, Pieter Abbeel, and Wojciech Zaremba (2017). "Hindsight Experience Replay". In: *Advances in Neural Information Processing Systems*. Ed. by I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett. Curran Associates, Inc., pp. 5048–5058. URL: http://papers.nips.cc/paper/7090-hindsight-experience-replay.pdf.

Asada, Minoru, Shoichi Noda, Sukoya Tawaratsumida, and Koh Hosoda (1996). "Purposive Behavior Acquisition for a Real Robot by Vision-Based Reinforcement Learning". In: *Machine Learning* 23.2, pp. 279–303. ISSN: 1573-0565. DOI: 10.1023/A:1018237008823.

Beeson, Patrick and Barrett Ames (2015). "TRAC-IK: An open-source library for improved solving of generic inverse kinematics". In: *IEEE-RAS International Conference on Humanoid Robots (Humanoids)*, pp. 928–935.

Literature II

Char, Ian, Youngseog Chung, Willie Neiswanger, Kirthevasan Kandasamy, Andrew Oakleigh Nelson, Mark Boyer, Egemen Kolemen, and Jeff Schneider (2019). "Offline Contextual Bayesian Optimization". In: *Advances in Neural Information Processing Systems*. Ed. by H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. Fox, and R. Garnett. Curran Associates, Inc., pp. 4627–4638.

Fabisch, Alexander (2019). "Empirical Evaluation of Contextual Policy Search with a Comparison-based Surrogate Model and Active Covariance Matrix Adaptation". In: *Genetic and Evolutionary Computation Conference Companion*. Ed. by Manuel López-Ibáñez. GECCO '19. ACM, pp. 251–252. ISBN: 978-1-4503-6748-6. DOI: 10.1145/3319619.3321935.

Florensa, Carlos, David Held, Xinyang Geng, and Pieter Abbeel (2018). "Automatic Goal Generation for Reinforcement Learning Agents". In: *International Conference on Machine Learning (ICML)*. Ed. by Jennifer Dy and Andreas Krause. Vol. 80. Proceedings of Machine Learning Research. Stockholmsmässan, Stockholm Sweden: PMLR, pp. 1515–1528.

Forestier, Sébastien and Pierre-Yves Oudeyer (2016). "Modular active curiosity-driven discovery of tool use". In: *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. Ed. by Il Hong Suh and Dong-Soo Kwon, pp. 3965–3972.

Graybiel, Ann M. (1998). "The basal ganglia and chunking of action repertoires". In: *Neurobiology of Learning and Memory* 70 (1), pp. 119–136. ISSN: 1074-7427. DOI: 10.1006/nlme.1998.3843.

Hansen, Jonas (2015). "Contextual Policy Search for Ball-Throwing on a Real Robot". MA thesis. Bremen, Germany: University of Bremen.

Literature III

Hansen, Nikolaus and Andreas Ostermeier (2001). "Completely Derandomized Self-Adaptation in Evolution Strategies". In: *Evolutionary Computation* 9.2, pp. 159–195. ISSN: 1063-6560. DOI: 10.1162/106365601750190398.

Kober, Jens, J. Andrew Bagnell, and Jan Peters (2013). "Reinforcement Learning in Robotics: A Survey". In: *International Journal of Robotics Research* 32.11, pp. 1238–1274. ISSN: 0278-3649. DOI: 10.1177/0278364913495721.

Kormushev, Petar, Sylvain Calinon, and Darwin G. Caldwell (2010). "Robot motor skill coordination with EM-based Reinforcement Learning". In: *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. Ed. by Ren C. Luo and Huei-Yung Lin, pp. 3232–3237. DOI: 10.1109/IROS.2010.5649089.

Levitis, Daniel A., William Z. Lidicker, and Glenn Freund (2009). "Behavioural biologists do not agree on what constitutes behaviour". In: *Animal Behaviour* 78.1, pp. 103–110. ISSN: 0003-3472. DOI: 10.1016/j.anbehav.2009.03.018.

Lillicrap, Timothy P., Jonathan J. Hunt, Alexander Pritzel, Nicolas Heess, Tom Erez, Yuval Tassa, David Silver, and Daan Wierstra (2016). "Continuous control with deep reinforcement learning". In: *International Conference on Learning Representations (ICLR)*. Ed. by Samy Bengio and Brian Kingsbury.

Mason, Matthew T. (2012). "Creation Myths: The Beginnings of Robotics Research". In: *IEEE Robotics Automation Magazine* 19.2, pp. 72–77. ISSN: 1070-9932. DOI: 10.1109/MRA.2012.2191437.

Literature IV

Metzen, Jan Hendrik (2015). "Active Contextual Entropy Search". In: *Workshop on Bayesian Optimization, Advances in Neural Information Processing Systems*. Ed. by Nando de Freitas, Ryan P. Adams, Bobak Shahriari, Roberto Calandra, and Amar Shah. Montreal, Quebec, Canada. URL: http://arxiv.org/abs/1511.04211.

Metzen, Jan Hendrik, Alexander Fabisch, and Jonas Hansen (2015). "Bayesian Optimization for Contextual Policy Search". In: *Machine Learning in Planning and Control of Robot Motion (MLPC) Workshop, IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. Ed. by Aleksandra Faust. URL: https://www.cs.unm.edu/~afaust/MLPC15_proceedings/MLPC15_paper_Metzen.pdf.

Mülling, Katharina, Jens Kober, Oliver Krömer, and Jan Peters (2013). "Learning to Select and Generalize Striking Movements in Robot Table Tennis". In: *International Journal of Robotics Research* 32.3.

Nocedal, Jorge (1980). "Updating quasi-Newton matrices with limited storage". In: *Mathematics of Computation* 35 (151), pp. 773–782. DOI: 10.2307/2006193.

OpenAI, Marcin Andrychowicz, Bowen Baker, Maciek Chociej, Rafal Józefowicz, Bob McGrew, Jakub Pachocki, Arthur Petron, Matthias Plappert, Glenn Powell, Alex Ray, Jonas Schneider, Szymon Sidor, Josh Tobin, Peter Welinder, Lilian Weng, and Wojciech Zaremba (2020). "Learning dexterous in-hand manipulation". In: *International Journal of Robotics Research* 39.1, pp. 3–20. DOI: 10.1177/0278364919887447.

Pearl, Judea (1983). *Heuristics: Intelligent Search Strategies for Computer Problem Solving*. Addison-Wesley. ISBN: 978-0-201-05594-8.

Literature V

Peng, Xue Bin, Pieter Abbeel, Sergey Levine, and Michiel van de Panne (2018). "DeepMimic: Example-Guided Deep Reinforcement Learning of Physics-Based Character Skills". In: *ACM Transactions on Graphics* 37.4. ISSN: 0730-0301. DOI: 10.1145/3197517.3201311.

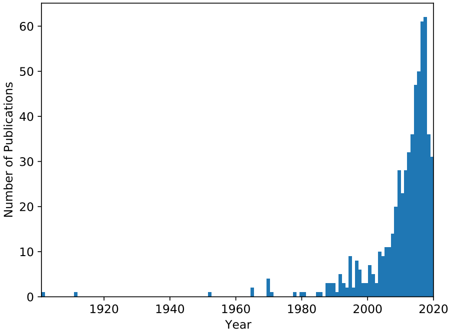
Pinsler, Robert, Peter Karkus, Andras Gabor Kupcsik, David Hsu, and Wee Sun Lee (2019). "Factored Contextual Policy Search with Bayesian optimization". In: *IEEE International Conference on Robotics and Automation (ICRA)*. Ed. by Ayanna Howard, pp. 7242–7248.

Rakita, Daniel, Bilge Mutlu, and Michael Gleicher (2017). "A Motion Retargeting Method for Effective Mimicry-Based Teleoperation of Robot Arms". In: *ACM/IEEE International Conference on Human-Robot Interaction*. Ed. by Astrid Weiss and James Young. HRI '17. Vienna, Austria: Association for Computing Machinery, pp. 361–370. ISBN: 9781450343367. DOI: 10.1145/2909824.3020254.

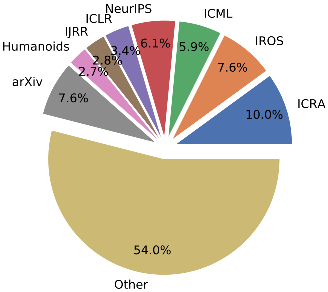
Scherzinger, Stefan, Arne Roennau, and Rüdiger Dillmann (2019). "Contact Skill Imitation Learning for Robot-Independent Assembly Programming". In: *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. Ed. by Dong Sun and Fumihito Arai, pp. 4309–4316.

Thrun, Sebastian and Tom M. Mitchell (1995). "Lifelong robot learning". In: *Robotics and Autonomous Systems* 15.1, pp. 25–46. ISSN: 0921-8890. DOI: 10.1016/0921-8890(95)00004-Y.

Literature



60
50
40
30
20
10
0
Number of Publications
1920
1940
1960
1980
2000
2020
Year



NeurIPS
ICML
ICLR
IJRR
Humanoids
IROS
3.4%
6.1%
5.9%
2.8%
7.6%
2.7%
arXiv
7.6%
ICRA
10.0%
54.0%
Other