A Tutorial on Meta-Reinforcement Learning

Other titles in Foundations and Trends® in Machine Learning

Generalization Bounds: Perspectives from Information Theory and PAC-Bayes
Fredrik Hellström, Giuseppe Durisi, Benjamin Guedj and Maxim Raginsky
ISBN: 978-1-63828-420-8

An Introduction to Deep Survival Analysis Models for Predicting Time-to-Event Outcomes
George H. Chen
ISBN: 978-1-63828-454-3

Automated Deep Learning: Neural Architecture Search Is Not the End
Xuanyi Dong, David Jacob Kedziora, Katarzyna Musial and Bogdan Gabrys
ISBN: 978-1-63828-318-8

AutonoML: Towards an Integrated Framework for Autonomous Machine Learning
David Jacob Kedziora, Katarzyna Musial and Bogdan Gabrys
ISBN: 978-1-63828-316-4

Causal Fairness Analysis: A Causal Toolkit for Fair Machine Learning
Drago Plečko and Elias Bareinboim
ISBN: 978-1-63828-330-0

User-friendly Introduction to PAC-Bayes Bounds
Pierre Alquier
ISBN: 978-1-63828-326-3

A Tutorial on Meta-Reinforcement Learning

Jacob Beck
University of Oxford
jacob.beck@cs.ox.ac.uk

Risto Vuorio
University of Oxford
risto.vuorio@cs.ox.ac.uk

Evan Zheran Liu
Stanford University
evanliu@cs.stanford.edu

Zheng Xiong
University of Oxford
zheng.xiong@cs.ox.ac.uk

Luisa Zintgraf
University of Oxford
zintgraf@deepmind.com

Chelsea Finn
Stanford University
cbfinn@cs.stanford.edu

Shimon Whiteson
University of Oxford
shimon.whiteson@cs.ox.ac.uk

now
the essence of knowledge
Boston — Delft

Foundations and Trends® in Machine Learning

Published, sold and distributed by:
now Publishers Inc.
PO Box 1024
Hanover, MA 02339
United States
Tel. +1-781-985-4510
www.nowpublishers.com
sales@nowpublishers.com

Outside North America:
now Publishers Inc.
PO Box 179
2600 AD Delft
The Netherlands
Tel. +31-6-51115274

The preferred citation for this publication is

J. Beck *et al.*. *A Tutorial on Meta-Reinforcement Learning*. Foundations and Trends® in Machine Learning, vol. 18, no. 2-3, pp. 224–384, 2025.

ISBN: 978-1-63828-541-0

Foundations and Trends® in Machine Learning
Volume 18, Issue 2-3, 2025

Editorial Scope

Foundations and Trends® welcomes monographs that touch on fundamental problems in machine learning from theoretical, methodological, and/or computational perspectives. We are particularly interested in monographs that seek to bridge such problems and perspectives with those from related fields, including (but not limited to) statistics, economics, and optimization.

Information for Librarians

Foundations and Trends® in Machine Learning, 2025, Volume 18, 6 issues. ISSN paper version 1935-8237. ISSN online version 1935-8245. Also available as a combined paper and online subscription.

Contents

A Tutorial on Meta-Reinforcement Learning

Jacob Beck[1*†], Risto Vuorio[1*‡], Evan Zheran Liu[2◊], Zheng Xiong[1], Luisa Zintgraf[1§], Chelsea Finn[2] and Shimon Whiteson[1]

[1] *University of Oxford, UK; jacob.beck@cs.ox.ac.uk, risto.vuorio@cs.ox.ac.uk, zheng.xiong@cs.ox.ac.uk, zintgraf@deepmind.com, shimon.whiteson@cs.ox.ac.uk*
[2] *Stanford University, USA; evanliu@cs.stanford.edu, cbfinn@cs.stanford.edu*

ABSTRACT

While deep reinforcement learning (RL) has fueled multiple high-profile successes in machine learning, it is held back from more widespread adoption by its often poor data efficiency and the limited generality of the policies it produces. A promising approach for alleviating these limitations is to cast the development of better RL algorithms as a machine learning problem itself in a process called meta-RL. Meta-RL is most commonly studied in a problem setting where, given a distribution of tasks, the goal is to learn a policy that is capable of adapting to any new task from the task distribution with as little data as possible. In this survey, we describe the meta-RL problem setting in detail as well as its major variations. We discuss how, at a high level, meta-RL

---

[*] Contributed equally.
[†] Now at Oracle.
[‡] Now at Inflection.
[◊] Now at Imbue.
[§] Now at DeepMind.

Jacob Beck, Risto Vuorio, Evan Zheran Liu, Zheng Xiong, Luisa Zintgraf, Chelsea Finn and Shimon Whiteson (2025), "A Tutorial on Meta-Reinforcement Learning", Foundations and Trends® in Machine Learning: Vol. 18, No. 2-3, pp 224–384. DOI: 10.1561/2200000080.

research can be clustered based on the presence of a task distribution and the learning budget available for each individual task. Using these clusters, we then survey meta-RL algorithms and applications. We conclude by presenting the open problems on the path to making meta-RL part of the standard toolbox for a deep RL practitioner.

1

Introduction

Meta-reinforcement learning (meta-RL) considers a family of machine learning (ML) methods that *learn to reinforcement learn*. That is, meta-RL methods use sample-inefficient ML to learn sample-efficient RL algorithms, or components thereof. As such, meta-RL is a special case of meta-learning (Vanschoren, 2018; Hospedales *et al.*, 2020; Huisman *et al.*, 2021), with the property that the learned algorithm is an RL algorithm. Meta-RL has been investigated as a machine learning problem for a significant period of time (Schmidhuber, 1987; Schmidhuber *et al.*, 1997; Thrun and Pratt, 1998; Schmidhuber, 2007). Intriguingly, research has also shown an analogue of meta-RL in the brain (Wang *et al.*, 2018).

Meta-RL has the potential to overcome some limitations of existing human-designed RL algorithms. While there has been significant progress in deep RL over the last several years, with success stories such as mastering the game of Go (Silver *et al.*, 2016), stratospheric balloon navigation (Bellemare *et al.*, 2020), or robot locomotion in challenging terrain (Miki *et al.*, 2022). RL remains highly sample inefficient, which limits its real-world applications. Meta-RL can produce (components of) RL algorithms that are much more sample efficient than existing RL methods, or even provide solutions to previously intractable problems.

At the same time, the promise of improved sample efficiency comes with two costs. First, meta-learning requires significantly more data than standard learning, as it trains an entire learning algorithm (often across multiple tasks). Second, meta-learning fits a learning algorithm to meta-training data, which may reduce its ability to generalize to other data. The trade-off that meta-learning offers is thus improved sample efficiency at test time, at the expense of sample efficiency during training and generality at test time.

Example application Consider, as a conceptual example, the task of automated cooking with a robot chef. When such a robot is deployed in somebody's kitchen, it must learn a kitchen-specific policy, since each kitchen has a different layout and appliances. This challenge is compounded by the fact that not all items needed for cooking are in plain sight; pots might be tucked away in cabinets, spices could be stored on high shelves, and utensils might be hidden in drawers. Therefore, the robot needs not only to understand the general layout but also remember where specific items are once discovered. Training the robot directly in a new kitchen from scratch is too time consuming and potentially dangerous due to random behavior early in training. One alternative is to pre-train the robot in a *single* training kitchen and then fine-tune it in the new kitchen. However, this approach does not take into account the subsequent fine-tuning procedure. In contrast, meta-RL would train the robot on a *distribution* of training kitchens such that it can adapt to any new kitchen in that distribution. This may entail learning some parameters to enable better fine-tuning, or learning the entire RL algorithm that will be deployed in the new kitchen. A robot trained this way can both make better use of the data collected and also collect better data, e.g., by focusing on the unusual or challenging features of the new kitchen. This meta-learning procedure requires more samples than the simple fine-tuning approach, but it only needs to occur once, and the resulting adaptation procedure can be significantly more sample efficient when deployed in the new test kitchen.

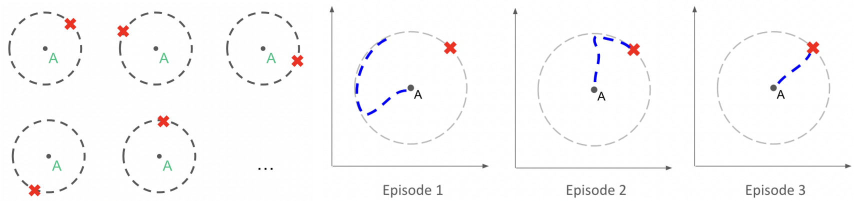
This example illustrates how, in general, meta-RL may be particularly useful when the need for efficient adaptation is frequent, so the cost of meta-training is relatively small. This includes, but is not limited to, safety-critical RL domains, where efficient data collection is necessary and exploration of novel behaviors is prohibitively costly or dangerous. In many cases, a large investment of sample-inefficient learning upfront (either with oversight, in a laboratory, or in simulation) is worthwhile to enable subsequent improved adaptation behavior. This example represents an aspirational application for meta-RL. In practice, meta-RL is applied to more limited robotics tasks such as robotic manipulation (Akkaya *et al.*, 2019; Zhao *et al.*, 2022b) and locomotion (Song *et al.*, 2020b).

Survey scope This is a survey of the meta-RL topic in machine learning and leaves out research on meta-RL in other fields such as neuroscience. Research on closely related machine learning topics is discussed in Section 2.6. To capture the breadth and depth of machine learning research on meta-RL, we surveyed the proceedings of several major machine learning conferences, as well as specialized workshops from the time period between 2017 and 2022. We found that a major portion of the meta-RL literature emerged post-2016, with the lion's share of contributions being concentrated in three conferences: NeurIPS, ICML, and ICLR. For a full list of conferences and workshops covered, see Appendix A. While our survey primarily emphasizes these conferences and the specified timeframe, we also discuss a selection of relevant papers from outside this scope. From the proceedings of these conferences and workshops, we searched for papers that explicitly mention meta-RL as well those that do not make an explicit reference but that we judged to nevertheless fit the topic. Finally, we do not claim exhaustive coverage of meta-RL research included in our survey scope but rather a holistic overview of the most salient ideas and general directions.

Survey overview The aim of this survey is to provide an entry point to meta-RL, as well as reflect on the current state of the field and open areas of research. In Section 2, we define meta-RL and the different problem settings it can be applied to, together with two example algorithms.

In Section 3, we consider the most prevalent problem setting in meta-RL: few-shot meta-RL. Here, the goal is to learn an RL algorithm capable of *fast adaptation*, i.e., learning a task within just a handful of episodes.

These algorithms are often trained on a given task distribution, and meta-learn how to efficiently adapt to any task from that distribution. A toy example to illustrate this setting is shown in Figure 1.1. Here, an agent is meta trained to learn how to navigate to different (initially unknown) goal positions on a 2D plane. At meta-test time, this agent can adapt efficiently to new tasks with unknown goal positions.



(a) Meta-Training Tasks **(b)** Rollout at Meta-Test Time

Figure 1.1: Example of the *fast adaptation* meta-RL problem setting discussed in Section 3. The agent (A) is meta-trained on a distribution of meta-training tasks to *learn* to go to goal position (X) located on a unit circle around its starting position **(a)**. At meta-test time, the agent can adapt quickly (within a handful of episodes) to new tasks with initially unknown goal positions **(b)**. In contrast, a standard RL algorithm may need hundreds of thousands of environment interactions when trained from scratch on one such task.

In Section 4, we consider many-shot settings. The goal here is to learn general-purpose RL algorithms not specific to a narrow task distribution, similar to those currently used in practice. There are two flavors of this: training on a distribution of tasks as above, or training on a single task but meta-learning alongside standard RL training.

Next, Section 5 presents some applications of meta-RL such as robotics. To conclude the survey, we discuss open problems in Section 6. These include generalization to broader task distributions for few-shot meta-RL, optimization challenges in many-shot meta-RL, and reduction of meta-training costs.

To provide high-level summaries of meta-RL research cited in this survey, we collect representative papers discussed in each section in a summary table presented within the section.

Appendix

A

List of Venues Surveyed

This survey is primarily based on the meta-RL research presented in the following conferences and workshops for the years from 2017 to 2022:

- International Conference on Learning Representations (ICLR)
- Conference on Neural Information Processing Systems (NeurIPS)
- International Conference on Machine Learning (ICML)
- Autonomous Agents and Multiagent Systems (AAMAS)
- Annual AAAI Conference on Artificial Intelligence (AAAI)
- Conference on Robot Learning (CoRL)
- Robotics: Science and Systems (RSS)
- International Conference on Intelligent Robots and Systems (IROS)
- NeurIPS Workshop on Meta-Learning
- ICLR Workshop on Meta-Learning

References

Agarwal, A., S. Kakade, A. Krishnamurthy, and W. Sun. (2020). "Flambe: Structural complexity and representation learning of low rank mdps". *Advances in neural information processing systems*. 33: 20095–20107.

Ajay, A., D. Ghosh, S. Levine, P. Agrawal, and A. Gupta. (2022). "Distributionally Adaptive Meta Reinforcement Learning". In: *Decision Awareness in Reinforcement Learning Workshop at ICML 2022*.

Akkaya, I., M. Andrychowicz, M. Chociej, M. Litwin, B. McGrew, A. Petron, A. Paino, M. Plappert, G. Powell, R. Ribas, *et al.* (2019). "Solving rubik's cube with a robot hand". *arXiv preprint arXiv:1910.07113*.

Akuzawa, K., Y. Iwasawa, and Y. Matsuo. (2021). "Estimating Disentangled Belief about Hidden State and Hidden Task for Meta-Reinforcement Learning". In: *Learning for Dynamics and Control*. PMLR. 73–86.

Alemi, A. A., I. Fischer, J. V. Dillon, and K. Murphy. (2017). "Deep Variational Information Bottleneck". In: *International Conference on Learning Representations*. URL: https://openreview.net/forum?id=HyxQzBceg.

Alet, F., M. F. Schneider, T. Lozano-Perez, and L. P. Kaelbling. (2020). "Meta-learning curiosity algorithms". In: *International Conference on Learning Representations*.

Almeida, D., C. Winter, J. Tang, and W. Zaremba. (2021). "A generalizable approach to learning optimizers". *arXiv preprint arXiv:2106.00958*.

Alver, S. and D. Precup. (2020). "A Brief Look at Generalization in Visual Meta-Reinforcement Learning". In: *4th Lifelong Machine Learning Workshop at ICML 2020*. URL: https://openreview.net/forum?id=WrCFtzVEwn.

Ammar, H. B., E. Eaton, M. E. Taylor, D. C. Mocanu, K. Driessens, G. Weiss, and K. Tuyls. (2014). "An automated measure of mdp similarity for transfer in reinforcement learning". In: *Workshops at the Twenty-Eighth AAAI Conference on Artificial Intelligence*.

Andrychowicz, M., M. Denil, S. Gomez, M. W. Hoffman, D. Pfau, T. Schaul, B. Shillingford, and N. De Freitas. (2016). "Learning to learn by gradient descent by gradient descent". *Advances in neural information processing systems*.

Anne, T., J. Wilkinson, and Z. Li. (2021). "Meta-Learning for Fast Adaptive Locomotion with Uncertainties in Environments and Robot Dynamics". In: *2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE. 4568–4575.

Arndt, K., M. Hazara, A. Ghadirzadeh, and V. Kyrki. (2020). "Meta Reinforcement Learning for Sim-to-real Domain Adaptation". In: *2020 IEEE International Conference on Robotics and Automation (ICRA)*. 2725–2731. DOI: 10.1109/ICRA40945.2020.9196540.

Arumugam, D. and S. Singh. (2022). "Planning to the Information Horizon of BAMDPs via Epistemic State Abstraction". In: *Advances in Neural Information Processing Systems*. Ed. by A. H. Oh, A. Agarwal, D. Belgrave, and K. Cho. URL: https://openreview.net/forum?id=7eUOC9fEIRO.

Aubret, A., L. Matignon, and S. Hassas. (2019). "A survey on intrinsic motivation in reinforcement learning". *arXiv preprint arXiv:1908.06976*.

Baxter, J. (2000). "A model of inductive bias learning". *Journal of artificial intelligence research*. 12: 149–198.

Bechtle, S., A. Molchanov, Y. Chebotar, E. Grefenstette, L. Righetti, G. Sukhatme, and F. Meier. (2021). "Meta Learning via Learned Loss". In: *2020 25th International Conference on Pattern Recognition (ICPR)*. IEEE Computer Society. 4161–4168.

Beck, J., K. Ciosek, S. Devlin, S. Tschiatschek, C. Zhang, and K. Hofmann. (2020). "AMRL: Aggregated Memory For Reinforcement Learning". In: *International Conference on Learning Representations*. URL: https://openreview.net/forum?id=Bkl7bREtDr.

Beck, J., M. Jackson, R. Vuorio, and S. Whiteson. (2022). "Hypernetworks in Meta-Reinforcement Learning". *CoRL*.

Beck, J., M. Jackson, R. Vuorio, Z. Xiong, and S. Whiteson. (2024). "SplAgger: Split Aggregation for Meta-Reinforcement Learning". *Reinforcement Learning Conference*.

Beck, J., S. Surana, M. McAuliffe, O. Bent, T. D. Barrett, J. J. G. Luis, and P. Duckworth. (2025). "Metalic: Meta-Learning In-Context with Protein Language Models". In: *The Thirteenth International Conference on Learning Representations*. URL: https://openreview.net/forum?id=TUKt7ag0qq.

Beck, J., R. Vuorio, Z. Xiong, and S. Whiteson. (2023). "Recurrent Hypernetworks are Surprisingly Strong in Meta-RL". In: *Thirty-seventh Conference on Neural Information Processing Systems*. URL: https://openreview.net/forum?id=pefAAzu8an.

Belkhale, S., R. Li, G. Kahn, R. McAllister, R. Calandra, and S. Levine. (2021). "Model-based meta-reinforcement learning for flight with suspended payloads". *IEEE Robotics and Automation Letters*. 6(2): 1471–1478.

Bellec, G., D. Salaj, A. Subramoney, R. Legenstein, and W. Maass. (2018). "Long short-term memory and Learning-to-learn in networks of spiking neurons". *NeurIPS*.

Bellemare, M. G., S. Candido, P. S. Castro, J. Gong, M. C. Machado, S. Moitra, S. S. Ponda, and Z. Wang. (2020). "Autonomous navigation of stratospheric balloons using reinforcement learning". *Nature*. 588(7836): 77–82.

Bellemare, M. G., Y. Naddaf, J. Veness, and M. Bowling. (2013). "The arcade learning environment: An evaluation platform for general agents". *Journal of Artificial Intelligence Research*. 47: 253–279.

Benjamins, C., T. Eimer, F. Schubert, A. Biedenkapp, B. Rosenhahn, F. Hutter, and M. Lindauer. (2021). "CARL: A Benchmark for Contextual and Adaptive Reinforcement Learning".

Berseth, G., Z. Zhang, G. Zhang, C. Finn, and S. Levine. (2021). "Comps: Continual meta policy search". *arXiv preprint arXiv:2112.04467*.

Bonardi, A., S. James, and A. J. Davison. (2020). "Learning One-Shot Imitation from Humans without Humans". *CVPR*.

Bonnet, C., P. Caron, T. D. Barrett, I. Davies, and A. Laterre. (2021). "One Step at a Time: Pros and Cons of Multi-Step Meta-Gradient Reinforcement Learning". *5th Workshop on Meta-Learning at NeurIPS*. URL: https://arxiv.org/abs/2111.00206.

Bonnet, C., L. I. Midgley, and A. Laterre. (2022). "Debiasing Meta-Gradient Reinforcement Learning by Learning the Outer Value Function". In: *Sixth Workshop on Meta-Learning at the Conference on Neural Information Processing Systems*. URL: https://openreview.net/forum?id=HFVmkIRJ4J.

Boutilier, C., C.-w. Hsu, B. Kveton, M. Mladenov, C. Szepesvari, and M. Zaheer. (2020). "Differentiable Meta-Learning of Bandit Policies". In: *Advances in Neural Information Processing Systems*. Ed. by H. Larochelle, M. Ranzato, R. Hadsell, M. Balcan, and H. Lin. Vol. 33. Curran Associates, Inc. 2122–2134. URL: https://proceedings.neurips.cc/paper/2020/file/171ae1bbb81475eb96287dd78565b38b-Paper.pdf.

Brockman, G., V. Cheung, L. Pettersson, J. Schneider, J. Schulman, J. Tang, and W. Zaremba. (2016). "Openai gym". *arXiv preprint arXiv:1606.01540*.

Brown, T., B. Mann, N. Ryder, M. Subbiah, J. D. Kaplan, P. Dhariwal, A. Neelakantan, P. Shyam, G. Sastry, A. Askell, S. Agarwal, A. Herbert-Voss, G. Krueger, T. Henighan, R. Child, A. Ramesh, D. Ziegler, J. Wu, C. Winter, C. Hesse, M. Chen, E. Sigler, M. Litwin, S. Gray, B. Chess, J. Clark, C. Berner, S. McCandlish, A. Radford, I. Sutskever, and D. Amodei. (2020). "Language Models are Few-Shot Learners". In: *Advances in Neural Information Processing Systems*.

Burda, Y., H. Edwards, A. Storkey, and O. Klimov. (2019). "Exploration by random network distillation". In: *International Conference on Learning Representations*. URL: https://openreview.net/forum?id=H1lJJnR5Ym.

Burega, B., J. D. Martin, and M. Bowling. (2022). "Learning to Prioritize Planning Updates in Model-based Reinforcement Learning". In: *Sixth Workshop on Meta-Learning at the Conference on Neural Information Processing Systems*. URL: https://openreview.net/forum?id=uR7ePjeB6z.

Chalvidal, M., T. Serre, and R. VanRullen. (2022). "Meta-Reinforcement Learning with Self-Modifying Networks". In: *Advances in Neural Information Processing Systems*. Ed. by A. H. Oh, A. Agarwal, D. Belgrave, and K. Cho. URL: https://openreview.net/forum?id=cYeYzaP-5AF.

Chan, S. C. Y., I. Dasgupta, J. Kim, D. Kumaran, A. K. Lampinen, and F. Hill. (2022). "Transformers generalize differently from information stored in context vs in weights". In: *Workshop on Memory in Artificial and Real Intelligence at NeurIPS*.

Charakorn, R., P. Manoonpong, and N. Dilokthanakul. (2021). "Learning to Cooperate with Unseen Agents Through Meta-Reinforcement Learning". In: *Proceedings of the 20th International Conference on Autonomous Agents and MultiAgent Systems*. 1478–1479.

Chen, C., H. Wei, N. Xu, G. Zheng, M. Yang, Y. Xiong, K. Xu, and Z. Li. (2020). "Toward a thousand lights: Decentralized deep reinforcement learning for large-scale traffic signal control". In: *Proceedings of the AAAI Conference on Artificial Intelligence*. Vol. 34. No. 04. 3414–3421.

Chen, X., J. Hu, C. Jin, L. Li, and L. Wang. (2022). "Understanding Domain Randomization for Sim-to-real Transfer". In: *International Conference on Learning Representations*. URL: https://openreview.net/forum?id=T8vZHIRTrY.

Chen, Y., M. W. Hoffman, S. G. Colmenarejo, M. Denil, T. P. Lillicrap, M. Botvinick, and N. Freitas. (2017). "Learning to learn without gradient descent by gradient descent". In: *International Conference on Machine Learning*. PMLR. 748–756.

Cheng, Y., S. Feng, J. Yang, H. Zhang, and Y. Liang. (2022). "Provable benefit of multitask representation learning in reinforcement learning". *Advances in Neural Information Processing Systems*. 35: 31741–31754.

Choshen, E. and A. Tamar. (2023). "ContraBAR: Contrastive Bayes-Adaptive Deep RL". *ICML*.

Chowdhery, A., S. Narang, J. Devlin, M. Bosma, G. Mishra, A. Roberts, P. Barham, H. W. Chung, C. Sutton, S. Gehrmann, P. Schuh, K. Shi, S. Tsvyashchenko, J. Maynez, A. Rao, P. Barnes, Y. Tay, N. Shazeer, V. Prabhakaran, E. Reif, N. Du, B. Hutchinson, R. Pope, J. Bradbury, J. Austin, M. Isard, G. Gur-Ari, P. Yin, T. Duke, A. Levskaya, S. Ghemawat, S. Dev, H. Michalewski, X. Garcia, V. Misra, K. Robinson, L. Fedus, D. Zhou, D. Ippolito, D. Luan, H. Lim, B. Zoph, A. Spiridonov, R. Sepassi, D. Dohan, S. Agrawal, M. Omernick, A. M. Dai, T. S. Pillai, M. Pellat, A. Lewkowycz, E. Moreira, R. Child, O. Polozov, K. Lee, Z. Zhou, X. Wang, B. Saeta, M. Diaz, O. Firat, M. Catasta, J. Wei, K. Meier-Hellstern, D. Eck, J. Dean, S. Petrov, and N. Fiedel. (2022). "PaLM: Scaling Language Modeling with Pathways".

Clavera, I., J. Rothfuss, J. Schulman, Y. Fujita, T. Asfour, and P. Abbeel. (2018). "Model-Based Reinforcement Learning via Meta-Policy Optimization". *CoRL*.

Cobbe, K., C. Hesse, J. Hilton, and J. Schulman. (2020). "Leveraging procedural generation to benchmark reinforcement learning". In: *International conference on machine learning*. PMLR. 2048–2056.

Collins, L., A. Mokhtari, S. Oh, and S. Shakkottai. (2022). "Maml and anil provably learn representations". In: *International Conference on Machine Learning*. PMLR. 4238–4310.

Cong, L., M. Görner, P. Ruppel, H. Liang, N. Hendrich, and J. Zhang. (2020). "Self-Adapting Recurrent Models for Object Pushing from Learning in Simulation". *IROS*.

Dance, C. R., J. Perez, and T. Cachet. (2021). "Demonstration-Conditioned Reinforcement Learning for Few-Shot Imitation". In: *Proceedings of the 38th International Conference on Machine Learning*. Ed. by M. Meila and T. Zhang. Vol. 139. *Proceedings of Machine Learning Research*. PMLR. 2376–2387. URL: https://proceedings.mlr.press/v139/dance21a.html.

Dasari, S. and A. Gupta. (2020). "Transformers for One-Shot Visual Imitation". *CoRL*. abs/2011.05970. URL: https://arxiv.org/abs/2011.05970.

Deleu, T. and Y. Bengio. (2018). "The effects of negative adaptation in model-agnostic meta-learning". *arXiv preprint arXiv:1812.02159*.

Denevi, G., C. Ciliberto, R. Grazzi, and M. Pontil. (2019). "Learning-to-learn stochastic gradient descent with biased regularization". In: *International Conference on Machine Learning*. PMLR. 1566–1575.

Dennis, M., N. Jaques, E. Vinitsky, A. Bayen, S. Russell, A. Critch, and S. Levine. (2020). "Emergent complexity and zero-shot transfer via unsupervised environment design". *Advances in neural information processing systems*. 33: 13049–13061.

Devlin, J., M.-W. Chang, K. Lee, and K. Toutanova. (2019). "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding". In: *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*.

Dorfman, R., I. Shenfeld, and A. Tamar. (2021). "Offline Meta Reinforcement Learning–Identifiability Challenges and Effective Data Collection Strategies". *Advances in Neural Information Processing Systems*. 34: 4607–4618.

Duan, Y., M. Andrychowicz, B. C. Stadie, J. Ho, J. Schneider, I. Sutskever, P. Abbeel, and W. Zaremba. (2017). "One-Shot Imitation Learning". In: *NIPS*.

Duan, Y., X. Chen, R. Houthooft, J. Schulman, and P. Abbeel. (2016a). "Benchmarking deep reinforcement learning for continuous control". In: *International conference on machine learning*. PMLR. 1329–1338.

Duan, Y., J. Schulman, X. Chen, P. L. Bartlett, I. Sutskever, and P. Abbeel. (2016b). "RL^2: Fast reinforcement learning via slow reinforcement learning". *arXiv preprint arXiv:1611.02779*.

Duff, M. O. and A. Barto. (2002). "Optimal Learning: Computational Procedures for Bayes-Adaptive Markov Decision Processes". *PhD thesis*.

Elawady, A., G. Chhablani, R. Ramrakhya, K. Yadav, D. Batra, Z. Kira, and A. Szot. (2024). "ReLIC: A Recipe for 64k Steps of In-Context Reinforcement Learning for Embodied AI". *arXiv*.

Ellis, B., S. Moalla, M. Samvelyan, M. Sun, A. Mahajan, J. N. Foerster, and S. Whiteson. (2022). "SMACv2: An Improved Benchmark for Cooperative Multi-Agent Reinforcement Learning". *arXiv preprint arXiv:2212.07489*.

Emukpere, D., X. Alameda-Pineda, and C. Reinke. (2021). "Successor Feature Neural Episodic Control". *arXiv preprint arXiv:2111.03110*.

Espeholt, L., H. Soyer, R. Munos, K. Simonyan, V. Mnih, T. Ward, Y. Doron, V. Firoiu, T. Harley, I. Dunning, *et al.* (2018). "Impala: Scalable distributed deep-rl with importance weighted actor-learner architectures". In: *International Conference on Machine Learning*. PMLR. 1407–1416.

Eysenbach, B., A. Gupta, J. Ibarz, and S. Levine. (2019). "Diversity is All You Need: Learning Skills without a Reward Function". *ICLR*.

Fakoor, R., P. Chaudhari, S. Soatto, and A. J. Smola. (2020). "Meta-Q-Learning". In: *International Conference on Learning Representations*.

Fallah, A., K. Georgiev, A. Mokhtari, and A. Ozdaglar. (2020). "Provably convergent policy gradient methods for model-agnostic meta-reinforcement learning". *arXiv preprint arXiv:2002.05135*.

Fallah, A., K. Georgiev, A. Mokhtari, and A. Ozdaglar. (2021). "On the convergence theory of debiased model-agnostic meta-reinforcement learning". *Advances in Neural Information Processing Systems*. 34: 3096–3107.

Farquhar, G., S. Whiteson, and J. Foerster. (2019). "Loaded DiCE: Trading off bias and variance in any-order score function gradient estimators for reinforcement learning". *Advances in Neural Information Processing Systems*.

Feng, X., O. Slumbers, Z. Wan, B. Liu, S. McAleer, Y. Wen, J. Wang, and Y. Yang. (2021). "Neural auto-curricula in two-player zero-sum games". *Advances in Neural Information Processing Systems*. 34.

Finn, C., P. Abbeel, and S. Levine. (2017a). "Model-agnostic meta-learning for fast adaptation of deep networks". In: *International Conference on Machine Learning*. PMLR. 1126–1135.

Finn, C. and S. Levine. (2018). "Meta-Learning and Universality: Deep Representations and Gradient Descent can Approximate any Learning Algorithm". *ICLR*.

Finn, C., A. Rajeswaran, S. Kakade, and S. Levine. (2019). "Online meta-learning". In: *International Conference on Machine Learning*. PMLR. 1920–1930.

Finn, C., T. Yu, T. Zhang, P. Abbeel, and S. Levine. (2017b). "One-Shot Visual Imitation Learning via Meta-Learning". *CoRL*.

Flennerhag, S., A. A. Rusu, R. Pascanu, F. Visin, H. Yin, and R. Hadsell. (2020). "Meta-Learning with Warped Gradient Descent". In: *International Conference on Learning Representations*. URL: https://openreview.net/forum?id=rkeiQlBFPB.

Flennerhag, S., Y. Schroecker, T. Zahavy, H. van Hasselt, D. Silver, and S. Singh. (2021). "Bootstrapped Meta-Learning". *arXiv preprint arXiv:2109.04504*.

Foerster, J., G. Farquhar, M. Al-Shedivat, T. Rocktäschel, E. Xing, and S. Whiteson. (2018a). "Dice: The infinitely differentiable monte carlo estimator". In: *International Conference on Machine Learning*. PMLR. 1529–1538.

Foerster, J. N., R. Y. Chen, M. Al-Shedivat, S. Whiteson, P. Abbeel, and I. Mordatch. (2018b). "Learning with Opponent-Learning Awareness". *AAMAS*.

Fortunato, M., M. Tan, R. Faulkner, S. Hansen, A. P. Badia, G. Buttimore, C. Deck, J. Z. Leibo, and C. Blundell. (2019). "Generalization of Reinforcement Learners with Working and Episodic Memory". *NeurIPS*.

Fu, H., H. Tang, J. Hao, C. Chen, X. Feng, D. Li, and W. Liu. (2021). "Towards effective context for meta-reinforcement learning: an approach based on contrastive learning". In: *Proceedings of the AAAI Conference on Artificial Intelligence*. Vol. 35. No. 8. 7457–7465.

Fujimoto, S., D. Meger, and D. Precup. (2019). "Off-Policy Deep Reinforcement Learning without Exploration". *ICML*.

Galashov, A., J. Schwarz, H. Kim, M. Garnelo, D. Saxton, P. Kohli, S. M. A. Eslami, and Y. W. Teh. (2019). "Meta-Learning surrogate models for sequential decision making".

Gao, C., Y. Jiang, and F. Chen. (2022). "Transferring Hierarchical Structures with Dual Meta Imitation Learning". In: *6th Annual Conference on Robot Learning*. URL: https://openreview.net/forum?id=MUcBYHjzqp7.

Gao, K. and O. Sener. (2020). "Modeling and Optimization Trade-off in Meta-learning". In: *Advances in Neural Information Processing Systems*. Ed. by H. Larochelle, M. Ranzato, R. Hadsell, M. Balcan, and H. Lin. Vol. 33. Curran Associates, Inc. 11154–11165. URL: https://proceedings.neurips.cc/paper/2020/file/7fc63ff01769c4fa7d9279e97e307829-Paper.pdf.

Gao, Y., E. Sibirtseva, G. Castellano, and D. Kragic. (2019). "Fast adaptation with meta-reinforcement learning for trust modelling in human-robot interaction". In: *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE. 305–312.

Garau-Luis, J. J., Y. Miao, J. D. Co-Reyes, A. Parisi, J. Tan, E. Real, and A. Faust. (2022). "Multi-objective evolution for generalizable policy gradient algorithms". In: *ICLR 2022 Workshop on Generalizable Policy Learning in Physical World*.

Garcia, F. M. and P. S. Thomas. (2019). "A Meta-MDP Approach to Exploration for Lifelong Reinforcement Learning". *NeurIPS*.

Garnelo, M., J. Schwarz, D. Rosenbaum, F. Viola, D. J. Rezende, S. M. A. Eslami, and Y. W. Teh. (2018). "Neural Processes". *ICML*.

Gerstgrasser, M. and D. C. Parkes. (2022). "Meta-RL for Multi-Agent RL: Learning to Adapt to Evolving Agents". In: *Sixth Workshop on Meta-Learning at the Conference on Neural Information Processing Systems*. URL: https://openreview.net/forum?id=0toY1f8-Iq9.

Ghadirzadeh, A., X. Chen, P. Poklukar, C. Finn, M. Björkman, and D. Kragic. (2021). "Bayesian Meta-Learning for Few-Shot Policy Adaptation Across Robotic Platforms". In: *2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. 1274–1280. DOI: 10.1109/IROS51168.2021.9636628.

Ghavamzadeh, M., S. Mannor, J. Pineau, A. Tamar, *et al.* (2015). "Bayesian reinforcement learning: A survey". *Foundations and Trends® in Machine Learning*. 8(5-6): 359–483.

Ghosh, D., A. Ajay, P. Agrawal, and S. Levine. (2022). "Offline rl policies should be trained to be adaptive". In: *International Conference on Machine Learning*. PMLR. 7513–7530.

Goldie, A. D., C. Lu, M. T. Jackson, S. Whiteson, and J. N. Foerster. (2024). "Can Learned Optimization Make Reinforcement Learning Less Difficult?" *Advances in Neural Information Processing Systems*.

Goo, W. and S. Niekum. (2019). "One-shot learning of multi-step tasks from observation via activity localization in auxiliary video". In: *2019 international conference on robotics and automation (ICRA)*. IEEE. 7755–7761.

Graves, A., G. Wayne, and I. Danihelka. (2014). "Neural Turing Machines". *arXiv*. abs/1410.5401. URL: http://arxiv.org/abs/1410.5401.

Greenberg, I., S. Mannor, G. Chechik, and E. Meirom. (2023). "Train Hard, Fight Easy: Robust Meta Reinforcement Learning". *arXiv preprint arXiv:2301.11147*.

Grewal, Y. S., F. de Nijs, and S. Goodwin. (2021). "Variance-Seeking Meta-Exploration to Handle Out-of-Distribution Tasks". In: *Deep RL Workshop NeurIPS 2021*.

Grigsby, J., L. Fan, and Y. Zhu. (2024). "AMAGO: Scalable In-Context Reinforcement Learning for Adaptive Agents". *ICLR*.

Grover, A., M. Al-Shedivat, J. Gupta, Y. Burda, and H. Edwards. (2018). "Learning Policy Representations in Multiagent Systems". In: *Proceedings of the 35th International Conference on Machine Learning*. Ed. by J. Dy and A. Krause. Vol. 80. *Proceedings of Machine Learning Research*. PMLR. 1802–1811. URL: https://proceedings.mlr.press/v80/grover18a.html.

Guez, A., D. Silver, and P. Dayan. (2013). "Scalable and efficient Bayes-adaptive reinforcement learning based on Monte-Carlo tree search". *Journal of Artificial Intelligence Research*. 48: 841–883.

Guo, Y., Q. Wu, and H. Lee. (2022). "Learning Action Translator for Meta Reinforcement Learning on Sparse-Reward Tasks". In: *Proceedings of the AAAI Conference on Artificial Intelligence*. Vol. 36. No. 6. 6792–6800.

Guo, Z. D., M. G. Azar, B. Piot, B. A. Pires, and R. Munos. (2018). "Neural predictive belief representations". *arXiv preprint arXiv:1811.06407*.

Gupta, A., M. Lanctot, and A. Lazaridou. (2021). "Dynamic population-based meta-learning for multi-agent communication with natural language". *NeurIPS*.

Gupta, A., B. Eysenbach, C. Finn, and S. Levine. (2018a). "Unsupervised Meta-Learning for Reinforcement Learning". *arXiv*. abs/1806.04640.

Gupta, A., R. Mendonca, Y. Liu, P. Abbeel, and S. Levine. (2018b). "Meta-Reinforcement Learning of Structured Exploration Strategies". In: *Advances in Neural Information Processing Systems*. Ed. by S. Bengio, H. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi, and R. Garnett. Vol. 31. Curran Associates, Inc. URL: https://proceedings.neurips.cc/paper/2018/file/4de754248c196c85ee4fbdcee89179bd-Paper.pdf.

Gurumurthy, S., S. Kumar, and K. Sycara. (2020). "MAME : Model-Agnostic Meta-Exploration". In: *Proceedings of the Conference on Robot Learning*. Ed. by L. P. Kaelbling, D. Kragic, and K. Sugiura. Vol. 100. *Proceedings of Machine Learning Research*. PMLR. 910–922. URL: https://proceedings.mlr.press/v100/gurumurthy20a.html.

Ha, D., A. Dai, and Q. V. Le. (2017). "Hypernetworks". In: *International Conference on Learning Representation (ICLR)*.

Haarnoja, T., A. Zhou, P. Abbeel, and S. Levine. (2018). "Soft actor-critic: Off-policy maximum entropy deep reinforcement learning with a stochastic actor". In: *International conference on machine learning*. PMLR. 1861–1870.

Hallak, A., D. D. Castro, and S. Mannor. (2015). "Contextual Markov Decision Processes".

Harrison, J., A. Sharma, R. Calandra, and M. Pavone. (2018). "Control adaptation via meta-learning dynamics". In: *Workshop on Meta-Learning at NeurIPS*. Vol. 2018.

Hausknecht, M. J. and P. Stone. (2015). "Deep Recurrent Q-Learning for Partially Observable MDPs". In: *AAAI Fall Symposia*.

He, J. Z.-Y., Z. Erickson, D. S. Brown, A. Raghunathan, and A. Dragan. (2022). "Learning Representations that Enable Generalization in Assistive Tasks". In: *6th Annual Conference on Robot Learning*. URL: https://openreview.net/forum?id=b88HF4vd_ej.

Hebb, D. O. (1949). *The organization of behavior; a neuropsychological theory*.

Heess, N., J. J. Hunt, T. P. Lillicrap, and D. Silver. (2015). "Memory-based control with recurrent neural networks". *NIPS Deep Reinforcement Learning Workshop*.

Hejna III, D. J. and D. Sadigh. (2022). "Few-Shot Preference Learning for Human-in-the-Loop RL". In: *6th Annual Conference on Robot Learning*. URL: https://openreview.net/forum?id=IKC5TfXLuW0.

Hessel, M., H. Soyer, L. Espeholt, W. Czarnecki, S. Schmitt, and H. Van Hasselt. (2019). "Multi-task deep reinforcement learning with popart". In: *Proceedings of the AAAI Conference on Artificial Intelligence*. Vol. 33. No. 01. 3796–3803.

Hiraoka, T., T. Imagawa, V. Tangkaratt, T. Osa, T. Onishi, and Y. Tsuruoka. (2021). "Meta-model-based meta-policy optimization". In: *Asian Conference on Machine Learning*. PMLR. 129–144.

Hochreiter, S., A. S. Younger, and P. R. Conwell. (2001). "Learning to learn using gradient descent". In: *International conference on artificial neural networks*. Springer. 87–94.

Hospedales, T., A. Antoniou, P. Micaelli, and A. Storkey. (2020). "Meta-learning in neural networks: A survey". *arXiv preprint arXiv:2004.05439*.

Houthooft, R., Y. Chen, P. Isola, B. Stadie, F. Wolski, O. Jonathan Ho, and P. Abbeel. (2018). "Evolved Policy Gradients". In: *Advances in Neural Information Processing Systems*. Ed. by S. Bengio, H. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi, and R. Garnett. Vol. 31. Curran Associates, Inc. URL: https://proceedings.neurips.cc/paper/2018/file/7876acb66640bad41f1e1371ef30c180-Paper.pdf.

Hu, J., X. Chen, C. Jin, L. Li, and L. Wang. (2021). "Near-optimal representation learning for linear bandits and linear rl". In: *International Conference on Machine Learning*. PMLR. 4349–4358.

Huisman, M., J. N. Van Rijn, and A. Plaat. (2021). "A survey of deep meta-learning". *Artificial Intelligence Review*. 54(6): 4483–4541.

Humplik, J., A. Galashov, L. Hasenclever, P. A. Ortega, Y. W. Teh, and N. Heess. (2019). "Meta reinforcement learning as task inference". *arXiv preprint arXiv:1905.06424*.

Imagawa, T., T. Hiraoka, and Y. Tsuruoka. (2022). "Off-Policy Meta-Reinforcement Learning With Belief-Based Task Inference". *IEEE Access*. 10: 49494–49507.

Jabri, A., K. Hsu, A. Gupta, B. Eysenbach, S. Levine, and C. Finn. (2019). "Unsupervised curricula for visual meta-reinforcement learning". *Advances in Neural Information Processing Systems*. 32.

Jackson, M. T., M. Jiang, J. Parker-Holder, R. Vuorio, C. Lu, G. Farquhar, S. Whiteson, and J. N. Foerster. (2023). "Discovering General Reinforcement Learning Algorithms with Adversarial Environment Design". *arXiv preprint arXiv:2310.02782*.

Jaderberg, M., V. Mnih, W. M. Czarnecki, T. Schaul, J. Z. Leibo, D. Silver, and K. Kavukcuoglu. (2016). "Reinforcement learning with unsupervised auxiliary tasks". *arXiv preprint arXiv:1611.05397*.

James, S., M. Bloesch, and A. J. Davison. (2018). "Task-Embedded Control Networks for Few-Shot Imitation Learning". In: *CoRL*.

James, S., Z. Ma, D. R. Arrojo, and A. J. Davison. (2020). "Rlbench: The robot learning benchmark & learning environment". *IEEE Robotics and Automation Letters*. 5(2): 3019–3026.

Jang, E., A. Irpan, M. Khansari, D. Kappler, F. Ebert, C. Lynch, S. Levine, and C. Finn. (2021). "BC-Z: Zero-Shot Task Generalization with Robotic Imitation Learning". In: *5th Annual Conference on Robot Learning*. URL: https://openreview.net/forum?id=8kbp23tSGYv.

Jin, C., Z. Yang, Z. Wang, and M. I. Jordan. (2020). "Provably efficient reinforcement learning with linear function approximation". In: *Conference on Learning Theory*. PMLR. 2137–2143.

Kaddour, J., S. Saemundsson, and M. Deisenroth (he/him). (2020). "Probabilistic Active Meta-Learning". In: *Advances in Neural Information Processing Systems*. Ed. by H. Larochelle, M. Ranzato, R. Hadsell, M. Balcan, and H. Lin. Vol. 33. Curran Associates, Inc. 20813–20822. URL: https://proceedings.neurips.cc/paper/2020/file/ef0d17b3bdb4ee2aa741ba28c7255c53-Paper.pdf.

Kamienny, P.-A., M. Pirotta, A. Lazaric, T. Lavril, N. Usunier, and L. Denoyer. (2020). "Learning adaptive exploration strategies in dynamic environments through informed policy regularization". *arXiv preprint arXiv:2005.02934*.

Katharopoulos, A., A. Vyas, N. Pappas, and F. Fleuret. (2020). "Transformers are RNNs: Fast Autoregressive Transformers with Linear Attention". *ICML*.

Kaushik, R., T. Anne, and J.-B. Mouret. (2020). "Fast online adaptation in robotics through meta-learning embeddings of simulated priors". In: *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE. 5269–5276.

Khodak, M., M.-F. F. Balcan, and A. S. Talwalkar. (2019). "Adaptive gradient-based meta-learning methods". *Advances in Neural Information Processing Systems*. 32.

Kim, D. K., M. Liu, M. D. Riemer, C. Sun, M. Abdulhai, G. Habibi, S. Lopez-Cot, G. Tesauro, and J. How. (2021). "A policy gradient algorithm for learning to learn in multiagent reinforcement learning". In: *International Conference on Machine Learning*. PMLR. 5541–5550.

Kirk, R., A. Zhang, E. Grefenstette, and T. Rocktäschel. (2023). "A survey of zero-shot generalisation in deep reinforcement learning". *Journal of Artificial Intelligence Research*. 76: 201–264.

Kirsch, L., S. Flennerhag, H. van Hasselt, A. Friesen, J. Oh, and Y. Chen. (2022a). "Introducing Symmetries to Black Box Meta Reinforcement Learning". *AAAI*.

Kirsch, L., J. Harrison, C. D. Freeman, J. Sohl-Dickstein, and J. Schmidhuber. (2023). "Towards General-Purpose In-Context Learning Agents". In: *NeurIPS 2023 Workshop on Distribution Shifts: New Frontiers with Foundation Models*.

Kirsch, L., J. Harrison, J. Sohl-Dickstein, and L. Metz. (2022b). "General-purpose in-context learning by meta-learning transformers". *arXiv*.

Kirsch, L., S. van Steenkiste, and J. Schmidhuber. (2019). "Improving generalization in meta reinforcement learning using learned objectives". *arXiv preprint arXiv:1910.04098*.

Korshunova, I., J. Degrave, J. Dambre, A. Gretton, and F. Huszar. (2020). "Exchangeable Models in Meta Reinforcement Learning". In: *4th Lifelong Machine Learning Workshop at ICML 2020*. URL: https://openreview.net/forum?id=TZFlzejFPT.

Kumar, A., Z. Fu, D. Pathak, and J. Malik. (2021a). "RMA: Rapid Motor Adaptation for Legged Robots". *RSS*. abs/2107.04034.

Kumar, S., I. Dasgupta, J. D. Cohen, N. D. Daw, and T. L. Griffiths. (2020). "Meta-Learning of Compositional Task Distributions in Humans and Machines". In: *4th Workshop on Meta-Learning at NeurIPS*.

Kumar, V. C. V., S. Ha, and C. K. Liu. (2021b). "Error-Aware Policy Learning: Zero-Shot Generalization in Partially Observable Dynamic Environments". In: *Proceedings of Robotics: Science and Systems*. Virtual. DOI: 10.15607/RSS.2021.XVII.065.

Küttler, H., N. Nardelli, A. Miller, R. Raileanu, M. Selvatici, E. Grefenstette, and T. Rocktäschel. (2020). "The nethack learning environment". *Advances in Neural Information Processing Systems*. 33: 7671–7684.

Kveton, B., M. Konobeev, M. Zaheer, C.-W. Hsu, M. Mladenov, C. Boutilier, and C. Szepesvari. (2021). "Meta-Thompson Sampling". In: *Proceedings of the 38th International Conference on Machine Learning*. Ed. by M. Meila and T. Zhang. Vol. 139. *Proceedings of Machine Learning Research*. PMLR. 5884–5893. URL: https://proceedings.mlr.press/v139/kveton21a.html.

Lan, L., Z. Li, X. Guan, and P. Wang. (2019). "Meta reinforcement learning with task embedding and shared policy". In: *Proceedings of the 28th International Joint Conference on Artificial Intelligence*. 2794–2800.

Lan, Q., A. R. Mahmood, S. Yan, and Z. Xu. (2023). "Learning to Optimize for Reinforcement Learning". *arXiv preprint arXiv:2302.01470*.

Lange, R. T. and H. Sprekeler. (2020). “Learning not to learn: Nature versus nurture in silico”. In: *4th Workshop on Meta-Learning at NeurIPS*.

Laskin, M., L. Wang, J. Oh, E. Parisotto, S. Spencer, R. Steigerwald, D. Strouse, S. Hansen, A. Filos, E. Brooks, *et al.* (2023). “In-context reinforcement learning with algorithm distillation”. *ICLR*.

Lee, B., J. Lee, P. Vrancx, D. Kim, and K.-E. Kim. (2020a). “Batch reinforcement learning with hyperparameter gradients”. In: *International Conference on Machine Learning*. PMLR. 5725–5735.

Lee, G., B. Hou, A. Mandalika, J. Lee, S. Choudhury, and S. S. Srinivasa. (2019). “Bayesian Policy Optimization for Model Uncertainty”. In: *International Conference on Learning Representations*. URL: https://openreview.net/forum?id=SJGvns0qK7.

Lee, J., A. Xie, A. Pacchiano, Y. Chandak, C. Finn, O. Nachum, and E. Brunskill. (2024). “Supervised pretraining can learn in-context reinforcement learning”. *Advances in Neural Information Processing Systems*. 36.

Lee, K., Y. Seo, S. Lee, H. Lee, and J. Shin. (2020b). “Context-aware Dynamics Model for Generalization in Model-Based Reinforcement Learning”. *ICML*.

Lee, S. and S.-Y. Chung. (2021). “Improving Generalization in Meta-RL with Imaginary Tasks from Latent Dynamics Mixture”. *Advances in Neural Information Processing Systems*. 34.

Levine, S., A. Kumar, G. Tucker, and J. Fu. (2020). “Offline reinforcement learning: Tutorial, review, and perspectives on open problems”. *arXiv preprint arXiv:2005.01643*.

Li, J., T. Lu, X. Cao, Y. Cai, and S. Wang. (2021a). “Meta-imitation learning by watching video demonstrations”. In: *International Conference on Learning Representations*.

Li, K. and J. Malik. (2016). “Learning to optimize”. *arXiv preprint arXiv:1606.01885*.

Li, K. and J. Malik. (2017). “Learning to Optimize”. *ICLR*.

Li, K., A. Gupta, A. Reddy, V. H. Pong, A. Zhou, J. Yu, and S. Levine. (2021b). "MURAL: Meta-Learning Uncertainty-Aware Rewards for Outcome-Driven Reinforcement Learning". In: *Proceedings of the 38th International Conference on Machine Learning*. Ed. by M. Meila and T. Zhang. Vol. 139. *Proceedings of Machine Learning Research*. PMLR. 6346–6356. URL: https://proceedings.mlr.press/v139/li21g.html.

Li, L., R. Yang, and D. Luo. (2021c). "FOCAL: Efficient Fully-Offline Meta-Reinforcement Learning via Distance Metric Learning and Behavior Regularization". In: *9th International Conference on Learning Representations, ICLR 2021, Virtual Event, Austria, May 3-7, 2021*. OpenReview.net. URL: https://openreview.net/forum?id=8cpHIfgY4Dj.

Li, Z., F. Zhou, F. Chen, and H. Li. (2017). "Meta-SGD: Learning to Learn Quickly for Few Shot Learning". *CoRR*. abs/1707.09835. URL: http://arxiv.org/abs/1707.09835.

Lillicrap, T. P., J. J. Hunt, A. Pritzel, N. Heess, T. Erez, Y. Tassa, D. Silver, and D. Wierstra. (2016). "Continuous control with deep reinforcement learning." In: *ICLR (Poster)*. URL: http://arxiv.org/abs/1509.02971.

Lin, S., J. Wan, T. Xu, Y. Liang, and J. Zhang. (2022). "Model-Based Offline Meta-Reinforcement Learning with Regularization". In: *International Conference on Learning Representations*. URL: https://openreview.net/forum?id=EBn0uInJZWh.

Lin, X., Harjatin, S. Baweja, G. Kantor, and D. Held. (2019). "Adaptive Auxiliary Task Weighting for Reinforcement Learning". In: *NeurIPS*.

Lin, Z., G. Thomas, G. Yang, and T. Ma. (2020a). "Model-based Adversarial Meta-Reinforcement Learning". In: *Advances in Neural Information Processing Systems*.

Lin, Z., G. Thomas, G. Yang, and T. Ma. (2020b). "Model-based adversarial meta-reinforcement learning". *Advances in Neural Information Processing Systems*. 33: 10161–10173.

Liotet, P., F. Vidaich, A. M. Metelli, and M. Restelli. (2022). "Lifelong hyper-policy optimization with multiple importance sampling regularization". In: *Proceedings of the AAAI Conference on Artificial Intelligence*. Vol. 36. No. 7. 7525–7533.

Liu, B., X. Feng, J. Ren, L. Mai, R. Zhu, H. Zhang, J. Wang, and Y. Yang. (2022a). "A Theoretical Understanding of Gradient Bias in Meta-Reinforcement Learning". In: *Advances in Neural Information Processing Systems*. Ed. by A. H. Oh, A. Agarwal, D. Belgrave, and K. Cho. URL: https://openreview.net/forum?id=p9zeOtKQXKs.

Liu, B., X. Feng, J. Ren, L. Mai, R. Zhu, H. Zhang, J. Wang, and Y. Yang. (2022b). "A theoretical understanding of gradient bias in meta-reinforcement learning". *Advances in Neural Information Processing Systems*. 35: 31059–31072.

Liu, E. Z., A. Raghunathan, P. Liang, and C. Finn. (2021). "Decoupling Exploration and Exploitation for Meta-Reinforcement Learning without Sacrifices". In: *International Conference on Machine Learning*. PMLR. 6925–6935.

Liu, E. Z., M. P. Stephan, A. Nie, C. J. Piech, E. Brunskill, and C. Finn. (2022c). "Giving Feedback on Interactive Student Programs with Meta-Exploration". In: *Advances in Neural Information Processing Systems*.

Liu, H., R. Socher, and C. Xiong. (2019). "Taming maml: Efficient unbiased meta-reinforcement learning". In: *International Conference on Machine Learning*. PMLR. 4061–4071.

Lou, B., N. Zhao, and J. Wang. (2021). "Meta-learning from sparse recovery". In: *Fifth Workshop on Meta-Learning at the Conference on Neural Information Processing Systems*.

Lu, C., J. Kuba, A. Letcher, L. Metz, C. Schroeder de Witt, and J. Foerster. (2022a). "Discovered Policy Optimisation". In: *Advances in Neural Information Processing Systems*. Ed. by S. Koyejo, S. Mohamed, A. Agarwal, D. Belgrave, K. Cho, and A. Oh. Vol. 35. Curran Associates, Inc. 16455–16468. URL: https://proceedings.neurips.cc/paper_files/paper/2022/file/688c7a82e31653e7c256c6c29fd3b438-Paper-Conference.pdf.

Lu, C., T. Willi, A. Letcher, and J. N. Foerster. (2022b). "Adversarial Cheap Talk". In: *Sixth Workshop on Meta-Learning at the Conference on Neural Information Processing Systems*. URL: https://openreview.net/forum?id=VE6FEZc0ppH.

Lu, C., T. Willi, C. A. S. De Witt, and J. Foerster. (2022c). "Model-free opponent shaping". In: *International Conference on Machine Learning*. PMLR. 14398–14411.

Luketina, J., S. Flennerhag, Y. Schroecker, D. Abel, T. Zahavy, and S. Singh. (2022). "Meta-Gradients in Non-Stationary Environments". In: *ICLR Workshop on Agent Learning in Open-Endedness*. URL: https://openreview.net/forum?id=SlzBXwZIZ9.

Luna Gutierrez, R. and M. Leonetti. (2020). "Information-theoretic task selection for meta-reinforcement learning". *Advances in Neural Information Processing Systems*. 33: 20532–20542.

Luo, F.-M., S. Jiang, Y. Yu, Z. Zhang, and Y.-F. Zhang. (2022). "Adapt to environment sudden changes by learning a context sensitive policy". In: *Proceedings of the AAAI Conference on Artificial Intelligence*. Vol. 36. No. 7. 7637–7646.

Mandi, Z., P. Abbeel, and S. James. (2022). "On the Effectiveness of Fine-tuning Versus Meta-RL for Robot Manipulation". In: *CoRL 2022 Workshop on Pre-training Robot Learning*.

Mao, J., J. Foerster, T. Rocktäschel, M. Al-Shedivat, G. Farquhar, and S. Whiteson. (2019). "A baseline for any order gradient estimation in stochastic computation graphs". In: *International Conference on Machine Learning*. PMLR. 4343–4351.

Mehta, B., T. Deleu, S. C. Raparthy, C. J. Pal, and L. Paull. (2020). "Curriculum in Gradient-Based Meta-Reinforcement Learning". In: *Beyond Tabula Rasa in RL Workshop at ICLR*.

Meier, R. and A. Mujika. (2022). "Open-Ended Reinforcement Learning with Neural Reward Functions". In: *Advances in Neural Information Processing Systems*. Ed. by A. H. Oh, A. Agarwal, D. Belgrave, and K. Cho. URL: https://openreview.net/forum?id=NL05_JGVg99.

Melo, L. C. (2022). "Transformers are Meta-Reinforcement Learners".

Mendonca, R., X. Geng, C. Finn, and S. Levine. (2020). "Meta-reinforcement learning robust to distributional shift via model identification and experience relabeling". *arXiv preprint arXiv:2006.07178*.

Mendonca, R., A. Gupta, R. Kralev, P. Abbeel, S. Levine, and C. Finn. (2019). "Guided Meta-Policy Search". *NeurIPS*.

Metz, L., C. D. Freeman, S. S. Schoenholz, and T. Kachman. (2021). "Gradients are not all you need". *arXiv preprint arXiv:2111.05803*.

Metz, L., N. Maheswaranathan, J. Nixon, D. Freeman, and J. Sohl-Dickstein. (2019). "Understanding and correcting pathologies in the training of learned optimizers". In: *International Conference on Machine Learning*. PMLR. 4556–4565.

Miconi, T., A. Rawal, J. Clune, and K. O. Stanley. (2019). "Backpropamine: training self-modifying neural networks with differentiable neuromodulated plasticity". In: *International Conference on Learning Representations*. URL: https://openreview.net/forum?id=r1lrAiA5Ym.

Miconi, T., K. Stanley, and J. Clune. (2018). "Differentiable plasticity: training plastic neural networks with backpropagation". In: *Proceedings of the 35th International Conference on Machine Learning*. Ed. by J. Dy and A. Krause. Vol. 80. *Proceedings of Machine Learning Research*. PMLR. 3559–3568. URL: https://proceedings.mlr.press/v80/miconi18a.html.

Miki, T., J. Lee, J. Hwangbo, L. Wellhausen, V. Koltun, and M. Hutter. (2022). "Learning robust perceptive locomotion for quadrupedal robots in the wild". *Science Robotics*. 7(62): eabk2822.

Mishra, N., M. Rohaninejad, X. Chen, and P. Abbeel. (2018). "A Simple Neural Attentive Meta-Learner". In: *International Conference on Learning Representations*. URL: https://openreview.net/forum?id=B1DmUzWAW.

Mitchell, E., R. Rafailov, X. B. Peng, S. Levine, and C. Finn. (2020). "Offline Meta-Reinforcement Learning with Advantage Weighting". *CoRR*. abs/2008.06043. URL: https://arxiv.org/abs/2008.06043.

Mitchell, E., R. Rafailov, X. B. Peng, S. Levine, and C. Finn. (2021). "Offline Meta-Reinforcement Learning with Advantage Weighting". In: *Proceedings of the 38th International Conference on Machine Learning*. Ed. by M. Meila and T. Zhang. Vol. 139. *Proceedings of Machine Learning Research*. PMLR. 7780–7791. URL: https://proceedings.mlr.press/v139/mitchell21a.html.

Mnih, V., A. P. Badia, M. Mirza, A. Graves, T. Lillicrap, T. Harley, D. Silver, and K. Kavukcuoglu. (2016). "Asynchronous methods for deep reinforcement learning". In: *International conference on machine learning*. PMLR. 1928–1937.

Moreno, P., J. Humplik, G. Papamakarios, B. A. Pires, L. Buesing, N. Heess, and T. Weber. (2018). "Neural belief states for partially observed domains". In: *NeurIPS 2018 workshop on reinforcement learning under partial observability*.

Mu, Y., Y. Zhuang, F. Ni, B. Wang, J. Chen, J. Hao, and P. Luo. (2022). "DOMINO: Decomposed Mutual Information Optimization for Generalized Context in Meta-Reinforcement Learning". In: *Advances in Neural Information Processing Systems*. Ed. by A. H. Oh, A. Agarwal, D. Belgrave, and K. Cho. URL: https://openreview.net/forum?id=CJGUABT_COm.

Mutti, M., M. Mancassola, and M. Restelli. (2022). "Unsupervised reinforcement learning in multiple environments". In: *Proceedings of the AAAI Conference on Artificial Intelligence*. Vol. 36. No. 7. 7850–7858.

Nagabandi, A., I. Clavera, S. Liu, R. S. Fearing, P. Abbeel, S. Levine, and C. Finn. (2019a). "Learning to Adapt in Dynamic, Real-World Environments through Meta-Reinforcement Learning". In: *International Conference on Learning Representations*.

Nagabandi, A., I. Clavera, S. Liu, R. S. Fearing, P. Abbeel, S. Levine, and C. Finn. (2019b). "Learning to Adapt in Dynamic, Real-World Environments through Meta-Reinforcement Learning". In: *International Conference on Learning Representations*. URL: https://openreview.net/forum?id=HyztsoC5Y7.

Nagabandi, A., C. Finn, and S. Levine. (2019c). "Deep Online Learning Via Meta-Learning: Continual Adaptation for Model-Based RL". In: *International Conference on Learning Representations*. URL: https://openreview.net/forum?id=HyxAfnA5tm.

Najarro, E. and S. Risi. (2020). "Meta-Learning through Hebbian Plasticity in Random Networks". In: *Advances in Neural Information Processing Systems*. Ed. by H. Larochelle, M. Ranzato, R. Hadsell, M. F. Balcan, and H. Lin. Vol. 33. Curran Associates, Inc. 20719–20731. URL: https://proceedings.neurips.cc/paper/2020/file/ee23e7ad9b473ad072d57aaa9b2a5222-Paper.pdf.

Nguyen, H. H., A. Baisero, D. Wang, C. Amato, and R. Platt. (2022). "Leveraging Fully Observable Policies for Learning under Partial Observability". In: *6th Annual Conference on Robot Learning*. URL: https://openreview.net/forum?id=pn-HOPBioUE.

Ni, T., B. Eysenbach, and R. Salakhutdinov. (2022). "Recurrent Model-Free RL Can Be a Strong Baseline for Many POMDPs". In: *Proceedings of the 39th International Conference on Machine Learning*.

Nichol, A., J. Achiam, and J. Schulman. (2018a). "On first-order meta-learning algorithms". *arXiv preprint arXiv:1803.02999*.

Nichol, A., V. Pfau, C. Hesse, O. Klimov, and J. Schulman. (2018b). "Gotta learn fast: A new benchmark for generalization in rl". *arXiv preprint arXiv:1804.03720*.

Oh, J., V. Chockalingam, S. P. Singh, and H. Lee. (2016). "Control of Memory, Active Perception, and Action in Minecraft". *ICML*.

Oh, J., M. Hessel, W. M. Czarnecki, Z. Xu, H. van Hasselt, S. Singh, and D. Silver. (2020). "Discovering reinforcement learning algorithms". *arXiv preprint arXiv:2007.08794*.

Packer, C., P. Abbeel, and J. E. Gonzalez. (2021). "Hindsight Task Relabelling: Experience Replay for Sparse Reward Meta-RL". In: *Advances in Neural Information Processing Systems*. URL: https://proceedings.neurips.cc/paper/2021/file/1454ca2270599546dfcd2a3700e4d2f1-Paper.pdf.

Paine, T. L., S. G. Colmenarejo, Z. Wang, S. Reed, Y. Aytar, T. Pfaff, M. W. Hoffman, G. Barth-Maron, S. Cabi, D. Budden, *et al.* (2018). "One-shot high-fidelity imitation: Training large-scale deep nets with rl". *arXiv preprint arXiv:1810.05017*.

Papoudakis, G. and S. V. Albrecht. (2020). "Variational Autoencoders for Opponent Modeling in Multi-Agent Systems". In: *AAAI 2020 Workshop on Reinforcement Learning in Games*.

Parisotto, E., H. F. Song, J. W. Rae, R. Pascanu, Ç. Gülçehre, S. M. Jayakumar, M. Jaderberg, R. L. Kaufman, A. Clark, S. Noury, M. M. Botvinick, N. Heess, and R. Hadsell. (2020). "Stabilizing Transformers for Reinforcement Learning". *ICML*.

Park, E. and J. B. Oliva. (2019). "Meta-Curvature". *NeurIPS*.

Park, S., O. Rybkin, and S. Levine. (2024). "METRA: Scalable Unsupervised RL with Metric-Aware Abstraction". In: *The Twelfth International Conference on Learning Representations*. URL: https://openreview.net/forum?id=c5pwL0Soay.

Parker-Holder, J., R. Rajan, X. Song, A. Biedenkapp, Y. Miao, T. Eimer, B. Zhang, V. Nguyen, R. Calandra, A. Faust, *et al.* (2022). "Automated Reinforcement Learning (AutoRL): A Survey and Open Problems". *arXiv preprint arXiv:2201.03916*.

Pascanu, R., T. Mikolov, and Y. Bengio. (2013). "On the difficulty of training recurrent neural networks". *ICML*.

Peng, M., B. Zhu, and J. Jiao. (2021). "Linear Representation Meta-Reinforcement Learning for Instant Adaptation". URL: https://openreview.net/forum?id=lNrtNGkr-vw.

Perez, C., F. P. Such, and T. Karaletsos. (2020). "Generalized hidden parameter mdps: Transferable model-based rl in a handful of trials". In: *Proceedings of the AAAI Conference on Artificial Intelligence*. Vol. 34. No. 04. 5403–5411.

Polydoros, A. S. and L. Nalpantidis. (2017). "Survey of model-based reinforcement learning: Applications on robotics". *Journal of Intelligent & Robotic Systems*. 86(2): 153–173.

Pong, V. H., A. V. Nair, L. M. Smith, C. Huang, and S. Levine. (2022). "Offline meta-reinforcement learning with online self-supervision". In: *International Conference on Machine Learning*. PMLR. 17811–17829.

Prat, A. and E. Johns. (2021). "PERIL: Probabilistic Embeddings for hybrid Meta-Reinforcement and Imitation Learning". URL: https://openreview.net/forum?id=BIIwfP55pp.

Rafailov, R., V. K. Vijay, T. Yu, A. Singh, M. Phielipp, and C. Finn. (2021). "The Reflective Explorer: Online Meta-Exploration from Offline Data in Realistic Robotic Tasks". In: *Deep RL Workshop NeurIPS 2021*.

Raghu, A., M. Raghu, S. Bengio, and O. Vinyals. (2020). "Rapid Learning or Feature Reuse? Towards Understanding the Effectiveness of MAML". *ICLR*.

Raileanu, R., M. Goldstein, A. Szlam, and R. Fergus. (2020). "Fast Adaptation via Policy-Dynamics Value Functions". *ICML*.

Rajendran, J., R. Lewis, V. Veeriah, H. Lee, and S. Singh. (2020). "How Should an Agent Practice?" *Proceedings of the AAAI Conference on Artificial Intelligence*. 34(04): 5454–5461. DOI: 10.1609/aaai.v34i04.5995.

Rakelly, K., A. Zhou, C. Finn, S. Levine, and D. Quillen. (2019). "Efficient off-policy meta-reinforcement learning via probabilistic context variables". In: *International conference on machine learning*. PMLR. 5331–5340.

Raparthy, S. C., E. Hambro, R. Kirk, M. Henaff, and R. Raileanu. (2023). "Generalization to New Sequential Decision Making Tasks with In-Context Learning". *NeurIPS 2023 Workshop on Foundation Models for Decision Making*.

Ravi, S. and H. Larochelle. (2017). "Optimization as a Model for Few-Shot Learning". In: *ICLR*.

Rechenberg, I. (1971). "Evolutionsstrategie: Optimierung technischer Systeme nach Prinzipien der biologischen Evolution. Dr.-Ing". *PhD thesis*. Thesis, Technical University of Berlin, Department of Process Engineering.

Ren, A. Z., B. Govil, T.-Y. Yang, K. R. Narasimhan, and A. Majumdar. (2022a). "Leveraging Language for Accelerated Learning of Tool Manipulation". In: *6th Annual Conference on Robot Learning*. URL: https://openreview.net/forum?id=nPw7jaGBrCG.

Ren, Z., A. Liu, Y. Liang, J. Peng, and J. Ma. (2022b). "Efficient Meta Reinforcement Learning for Preference-based Fast Adaptation". In: *Advances in Neural Information Processing Systems*. Ed. by A. H. Oh, A. Agarwal, D. Belgrave, and K. Cho. URL: https://openreview.net/forum?id=61UwgeIotn.

Rengarajan, D., S. Chaudhary, J. Kim, D. Kalathil, and S. Shakkottai. (2022). "Enhanced Meta Reinforcement Learning via Demonstrations in Sparse Reward Environments". In: *Advances in Neural Information Processing Systems*. Ed. by A. H. Oh, A. Agarwal, D. Belgrave, and K. Cho. URL: https://openreview.net/forum?id=kCtnkLv-_W0.

Co-Reyes, J. D., S. Feng, G. Berseth, J. Qui, and S. Levine. (2021a). "Accelerating Online Reinforcement Learning via Model-Based Meta-Learning". In: *Learning to Learn - Workshop at ICLR 2021*. URL: https://openreview.net/forum?id=XCJ5PEkuMkC.

Co-Reyes, J. D., Y. Miao, D. Peng, E. Real, Q. V. Le, S. Levine, H. Lee, and A. Faust. (2021b). "Evolving Reinforcement Learning Algorithms". In: *International Conference on Learning Representations*. URL: https://openreview.net/forum?id=0XXpJ4OtjW.

Riemer, M., I. Cases, R. Ajemian, M. Liu, I. Rish, Y. Tu, and G. Tesauro. (2019). "Learning to Learn without Forgetting By Maximizing Transfer and Minimizing Interference". In: *International Conference on Learning Representations*. URL: https://openreview.net/forum?id=B1gTShAct7.

Rimon, Z., A. Tamar, and G. Adler. (2022). "Meta Reinforcement Learning with Finite Training Tasks - a Density Estimation Approach". In: *Advances in Neural Information Processing Systems*. Ed. by A. H. Oh, A. Agarwal, D. Belgrave, and K. Cho. URL: https://openreview.net/forum?id=Y-sdZLIi9R9.

Ritter, S., R. Faulkner, L. Sartran, A. Santoro, M. Botvinick, and D. Raposo. (2021). "Rapid Task-Solving in Novel Environments". In: *International Conference on Learning Representations*. URL: https://openreview.net/forum?id=F-mvpFpn_0q.

Ritter, S., J. Wang, Z. Kurth-Nelson, S. Jayakumar, C. Blundell, R. Pascanu, and M. Botvinick. (2018a). "Been There, Done That: Meta-Learning with Episodic Recall". In: *ICML*.

Ritter, S., J. Wang, Z. Kurth-Nelson, S. Jayakumar, C. Blundell, R. Pascanu, and M. Botvinick. (2018b). "Been There, Done That: Meta-Learning with Episodic Recall". In: *Proceedings of the 35th International Conference on Machine Learning*. Ed. by J. Dy and A. Krause. Vol. 80. *Proceedings of Machine Learning Research*. PMLR. 4354–4363. URL: https://proceedings.mlr.press/v80/ritter18a.html.

Rohani, S. R. R., S. Hedayatian, and M. S. Baghshah. (2022). "BIMRL: Brain Inspired Meta Reinforcement Learning". In: *2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE. 9048–9053.

Rothfuss, J., D. Lee, I. Clavera, T. Asfour, and P. Abbeel. (2019). "ProMP: Proximal Meta-Policy Search". In: *International Conference on Learning Representations*. URL: https://openreview.net/forum?id=SkxXCi0qFX.

Sæmundsson, S., K. Hofmann, and M. P. Deisenroth. (2018). "Meta reinforcement learning with latent variable gaussian processes". *arXiv preprint arXiv:1803.07551*.

Salimans, T., J. Ho, X. Chen, S. Sidor, and I. Sutskever. (2017). "Evolution strategies as a scalable alternative to reinforcement learning". *arXiv preprint arXiv:1703.03864*.

Santoro, A., S. Bartunov, M. Botvinick, D. Wierstra, and T. Lillicrap. (2016). "Meta-learning with memory-augmented neural networks". In: *International conference on machine learning*. PMLR. 1842–1850.

Schmidhuber, J. (1987). "Evolutionary principles in self-referential learning, or on learning how to learn: the meta-meta-... hook". *PhD thesis*. Technische Universität München.

Schmidhuber, J. (2007). "Gödel machines: Fully self-referential optimal universal self-improvers". In: *Artificial general intelligence*. Springer. 199–226.

Schmidhuber, J., J. Zhao, and M. Wiering. (1997). "Shifting inductive bias with success-story algorithm, adaptive Levin search, and incremental self-improvement". *Machine Learning*. 28(1): 105–130.

Schmitt, S., M. Hessel, and K. Simonyan. (2020). "Off-policy actor-critic with shared experience replay". In: *International Conference on Machine Learning*. PMLR. 8545–8554.

Schoettler, G., A. Nair, J. A. Ojea, S. Levine, and E. Solowjow. (2020). "Meta-reinforcement learning for robotic industrial insertion tasks". In: *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE. 9728–9735.

Schulman, J., F. Wolski, P. Dhariwal, A. Radford, and O. Klimov. (2017). "Proximal policy optimization algorithms". *arXiv preprint arXiv:1707.06347*.

Schwartzwald, A. and N. Papanikolopoulos. (2020). "Sim-to-Real with Domain Randomization for Tumbling Robot Control". In: *IROS*.

Seo, Y., D. Hafner, H. Liu, F. Liu, S. James, K. Lee, and P. Abbeel. (2022). "Masked World Models for Visual Control". In: *6th Annual Conference on Robot Learning*. URL: https://openreview.net/forum?id=Bf6on28H0Jv.

Seyed Ghasemipour, S. K., S. (Gu, and R. Zemel. (2019). "SMILe: Scalable Meta Inverse Reinforcement Learning through Context-Conditional Policies". In: *Advances in Neural Information Processing Systems*. Ed. by H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. Fox, and R. Garnett. Vol. 32. Curran Associates, Inc. URL: https://proceedings.neurips.cc/paper/2019/file/2b8f621e9244cea5007bac8f5d50e476-Paper.pdf.

Shala, G., S. P. Arango, A. Biedenkapp, F. Hutter, and J. Grabocka. (2022). "AutoRL-Bench 1.0". In: *Sixth Workshop on Meta-Learning at the Conference on Neural Information Processing Systems*. URL: https://openreview.net/forum?id=RyAl60VhTcG.

Sharaf, A. and H. Daumé III. (2021). "Meta-Learning Effective Exploration Strategies for Contextual Bandits". *Proceedings of the AAAI Conference on Artificial Intelligence*: 9541–9548.

Al-Shedivat, M., T. Bansal, Y. Burda, I. Sutskever, I. Mordatch, and P. Abbeel. (2018a). "Continuous Adaptation via Meta-Learning in Nonstationary and Competitive Environments". In: *International Conference on Learning Representations*. URL: https://openreview.net/forum?id=Sk2u1g-0-.

Al-Shedivat, M., T. Bansal, Y. Burda, I. Sutskever, I. Mordatch, and P. Abbeel. (2018b). "Continuous Adaptation via Meta-Learning in Nonstationary and Competitive Environments". In: *International Conference on Learning Representations*.

Shin, J., A. Hakobyan, M. Park, Y. Kim, G. Kim, and I. Yang. (2022). "Infusing model predictive control into meta-reinforcement learning for mobile robots in dynamic environments". *IEEE Robotics and Automation Letters*. 7(4): 10065–10072.

Silver, D., A. Huang, C. J. Maddison, A. Guez, L. Sifre, G. Van Den Driessche, J. Schrittwieser, I. Antonoglou, V. Panneershelvam, M. Lanctot, *et al.* (2016). "Mastering the game of Go with deep neural networks and tree search". *nature*. 529(7587): 484–489.

Simchowitz, M., C. Tosh, A. Krishnamurthy, D. J. Hsu, T. Lykouris, M. Dudik, and R. E. Schapire. (2021). "Bayesian decision-making under misspecified priors with applications to meta-learning". *Advances in Neural Information Processing Systems*. 34: 26382–26394.

Singh, A., E. Jang, A. Irpan, D. Kappler, M. Dalal, S. Levinev, M. Khansari, and C. Finn. (2020). "Scalable multi-task imitation learning with autonomous improvement". In: *2020 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE. 2167–2173.

Singh, S., R. L. Lewis, and A. G. Barto. (2009). "Where do rewards come from". In: *Proceedings of the annual conference of the cognitive science society*. Cognitive Science Society. 2601–2606.

Snell, J., K. Swersky, and R. Zemel. (2017). "Prototypical Networks for Few-shot Learning". 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.

Sohn, S., H. Woo, J. Choi, and H. Lee. (2020). "Meta Reinforcement Learning with Autonomous Inference of Subtask Dependencies". In: *International Conference on Learning Representations*. URL: https://openreview.net/forum?id=HkgsWxrtPB.

Song, X., W. Gao, Y. Yang, K. Choromanski, A. Pacchiano, and Y. Tang. (2020a). "ES-MAML: Simple Hessian-Free Meta Learning". In: *International Conference on Learning Representations*. URL: https://openreview.net/forum?id=S1exA2NtDB.

Song, X., Y. Yang, K. Choromanski, K. Caluwaerts, W. Gao, C. Finn, and J. Tan. (2020b). "Rapidly Adaptable Legged Robots via Evolutionary Meta-Learning". *IROS*.

Stadie, B. C., G. Yang, R. Houthooft, X. Chen, Y. Duan, Y. Wu, P. Abbeel, and I. Sutskever. (2018). "Some Considerations on Learning to Explore via Meta-Reinforcement Learning". *NeurIPS*.

Staley, E. W., C. Ashcraft, B. Stoler, J. Markowitz, G. Vallabha, C. Ratto, and K. D. Katyal. (2021). "Meta Arcade: A Configurable Environment Suite for Meta-Learning". *arXiv preprint arXiv:2112.00583*.

Stone, P., G. A. Kaminka, S. Kraus, and J. S. Rosenschein. (2010). "Ad Hoc Autonomous Agent Teams: Collaboration without Pre-Coordination". In: *Proceedings of the Twenty-Fourth Conference on Artificial Intelligence*.

Subramanian, J., A. Sinha, R. Seraj, and A. Mahajan. (2022). "Approximate information state for approximate planning and reinforcement learning in partially observed systems". *The Journal of Machine Learning Research*. 23(1): 483–565.

Sung, F., Y. Yang, L. Zhang, T. Xiang, P. H. S. Torr, and T. M. Hospedales. (2018). "Learning to Compare: Relation Network for Few-Shot Learning". *CVPR*.

Sung, F., L. Zhang, T. Xiang, T. M. Hospedales, and Y. Yang. (2017). "Learning to Learn: Meta-Critic Networks for Sample Efficient Learning". *CoRR*. abs/1706.09529. arXiv: 1706.09529. URL: http://arxiv.org/abs/1706.09529.

Sutskever, I. (2013). *Training recurrent neural networks*. University of Toronto Toronto, ON, Canada.

Sutton, R. S. (1992). "Adapting bias by gradient descent: An incremental version of delta-bar-delta". In: *AAAI*. San Jose, CA. 171–176.

Tack, J., J. Park, H. Lee, J. Lee, and J. Shin. (2022). "Meta-Learning with Self-Improving Momentum Target". In: *Advances in Neural Information Processing Systems*. Ed. by A. H. Oh, A. Agarwal, D. Belgrave, and K. Cho. URL: https://openreview.net/forum?id=FCNMbF_TsKm.

Tamar, A., D. Soudry, and E. Zisselman. (2022). "Regularization guarantees generalization in bayesian reinforcement learning through algorithmic stability". In: *Proceedings of the AAAI Conference on Artificial Intelligence*. Vol. 36. No. 8. 8423–8431.

Tang, Y. (2022). "Biased Gradient Estimate with Drastic Variance Reduction for Meta Reinforcement Learning". In: *International Conference on Machine Learning*. PMLR. 21050–21075.

Tang, Y., T. Kozuno, M. Rowland, R. Munos, and M. Valko. (2021). "Unifying gradient estimators for meta-reinforcement learning via off-policy evaluation". *Advances in Neural Information Processing Systems*. 34: 5303–5315.

Team, A. A., J. Bauer, K. Baumli, S. Baveja, F. Behbahani, A. Bhoopchand, N. Bradley-Schmieg, M. Chang, N. Clay, A. Collister, *et al.* (2023). "Human-Timescale Adaptation in an Open-Ended Task Space". *arXiv preprint arXiv:2301.07608*.

Teh, Y., V. Bapst, W. M. Czarnecki, J. Quan, J. Kirkpatrick, R. Hadsell, N. Heess, and R. Pascanu. (2017). "Distral: Robust multitask reinforcement learning". *Advances in neural information processing systems*. 30.

Tessera, K.-a. A., A. Rahman, and S. V. Albrecht. (2024). "HyperMARL: Adaptive Hypernetworks for Multi-Agent RL".

Thompson, W. R. (1933). "On the likelihood that one unknown probability exceeds another in view of the evidence of two samples". *Biometrika*. 25(3-4): 285–294.

Thrun, S. and L. Pratt. (1998). "Learning to learn: Introduction and overview". In: *Learning to learn*. Springer. 3–17.

Todorov, E., T. Erez, and Y. Tassa. (2012). "Mujoco: A physics engine for model-based control". In: *2012 IEEE/RSJ international conference on intelligent robots and systems*. IEEE. 5026–5033.

Touati, A. and Y. Ollivier. (2021). "Learning one representation to optimize all rewards". *Advances in Neural Information Processing Systems*. 34: 13–23.

Triantafillou, E., T. Zhu, V. Dumoulin, P. Lamblin, U. Evci, K. Xu, R. Goroshin, C. Gelada, K. Swersky, P.-A. Manzagol, *et al.* (2020). "Meta-dataset: A dataset of datasets for learning to learn from few examples". *International Conference on Learning Representations*.

Tripuraneni, N., C. Jin, and M. Jordan. (2021). "Provable meta-learning of linear representations". In: *International Conference on Machine Learning*. PMLR. 10434–10443.

Valiant, L. G. (1984). "A theory of the learnable". *Communications of the ACM*. 27(11): 1134–1142.

Vanschoren, J. (2018). "Meta-learning: A survey". *arXiv preprint arXiv:1810.03548*.

Vaswani, A., N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, and I. Polosukhin. (2017). "Attention Is All You Need". *NeurIPS*.

Vázquez-Canteli, J. R., S. Dey, G. Henze, and Z. Nagy. (2020). "CityLearn: Standardizing research in multi-agent reinforcement learning for demand response and urban energy management". *arXiv preprint arXiv:2012.10504*.

Vázquez-Canteli, J. R. and Z. Nagy. (2019). “Reinforcement learning for demand response: A review of algorithms and modeling techniques”. *Applied energy*. 235: 1072–1089.

Veeriah, V., M. Hessel, Z. Xu, J. Rajendran, R. L. Lewis, J. Oh, H. P. van Hasselt, D. Silver, and S. Singh. (2019). “Discovery of Useful Questions as Auxiliary Tasks”. In: *Advances in Neural Information Processing Systems*. Ed. by H. Wallach, H. Larochelle, A. Beygelzimer, F. d’Alché-Buc, E. Fox, and R. Garnett. Vol. 32. Curran Associates, Inc. URL: https://proceedings.neurips.cc/paper/2019/file/10ff0b5e85e5b85cc3095d431d8c08b4-Paper.pdf.

Veeriah, V., T. Zahavy, M. Hessel, Z. Xu, J. Oh, I. Kemaev, H. P. van Hasselt, D. Silver, and S. Singh. (2021). “Discovery of Options via Meta-Learned Subgoals”. In: *Advances in Neural Information Processing Systems*. Ed. by M. Ranzato, A. Beygelzimer, Y. Dauphin, P. Liang, and J. W. Vaughan. Vol. 34. Curran Associates, Inc. 29861–29873. URL: https://proceedings.neurips.cc/paper/2021/file/fa246d0262c3925617b0c72bb20eeb1d-Paper.pdf.

Vinyals, O., C. Blundell, T. Lillicrap, k. kavukcuoglu koray, and D. Wierstra. (2016). “Matching Networks for One Shot Learning”. In: *Advances in Neural Information Processing Systems*.

Vuorio, R., J. A. Beck, G. Farquhar, J. N. Foerster, and S. Whiteson. (2021). “No DICE: An Investigation of the Bias-Variance Tradeoff in Meta-Gradients”. In: *Deep RL Workshop NeurIPS 2021*.

Vuorio, R., S.-H. Sun, H. Hu, and J. Lim. (2019). “Multimodal Model-Agnostic Meta-Learning via Task-Aware Modulation”. In: *NeurIPS*.

Walke, H. R., J. H. Yang, A. Yu, A. Kumar, J. Orbik, A. Singh, and S. Levine. (2022). “Don’t Start From Scratch: Leveraging Prior Data to Automate Robotic Reinforcement Learning”. In: *6th Annual Conference on Robot Learning*. URL: https://openreview.net/forum?id=WbdaYyDkNZL.

Wan, M., J. Peng, and T. Gangwani. (2022). “Hindsight Foresight Relabeling for Meta-Reinforcement Learning”. In: *International Conference on Learning Representations*. URL: https://openreview.net/forum?id=P7OVkHEoHOZ.

Wang, H., J. Zhou, and X. He. (2020a). "Learning Context-aware Task Reasoning for Efficient Meta Reinforcement Learning". In: *Proceedings of the 19th International Conference on Autonomous Agents and MultiAgent Systems*. 1440–1448.

Wang, J. X., Z. Kurth-Nelson, D. Kumaran, D. Tirumala, H. Soyer, J. Z. Leibo, D. Hassabis, and M. Botvinick. (2018). "Prefrontal cortex as a meta-reinforcement learning system". *Nature neuroscience*. 21(6): 860–868.

Wang, J. X., Z. Kurth-Nelson, D. Tirumala, H. Soyer, J. Z. Leibo, R. Munos, C. Blundell, D. Kumaran, and M. Botvinick. (2016). "Learning to reinforcement learn". *arXiv preprint arXiv:1611.05763*.

Wang, J. X., M. King, N. Porcel, Z. Kurth-Nelson, T. Zhu, C. Deck, P. Choy, M. Cassin, M. Reynolds, H. F. Song, G. Buttimore, D. P. Reichert, N. C. Rabinowitz, L. Matthey, D. Hassabis, A. Lerchner, and M. M. Botvinick. (2021). "Alchemy: A structured task distribution for meta-reinforcement learning". *NeurIPS*.

Wang, Q. and H. van Hoof. (2022). "Learning Expressive Meta-Representations with Mixture of Expert Neural Processes". In: *Advances in Neural Information Processing Systems*. Ed. by A. H. Oh, A. Agarwal, D. Belgrave, and K. Cho. URL: https://openreview.net/forum?id=ju38DG3sbg6.

Wang, Q., Y. Lv, Y. Feng, Z. Xie, and J. Huang. (2023). "A Simple Yet Effective Strategy to Robustify the Meta Learning Paradigm". *Advances in Neural Information Processing Systems*.

Wang, Q. and H. Van Hoof. (2022). "Model-based Meta Reinforcement Learning using Graph Structured Surrogate Models and Amortized Policy Search". In: *Proceedings of the 39th International Conference on Machine Learning*.

Wang, Y., Q. Yao, J. T. Kwok, and L. M. Ni. (2020b). "Generalizing from a few examples: A survey on few-shot learning". *ACM computing surveys (csur)*. 53(3): 1–34.

Wang, Z., Y. Zhao, P. Yu, R. Zhang, and C. Chen. (2020c). "Bayesian Meta Sampling for Fast Uncertainty Adaptation". In: *International Conference on Learning Representations*. URL: https://openreview.net/forum?id=Bkxv90EKPB.

Weihs, L., U. Jain, I.-J. Liu, J. Salvador, S. Lazebnik, A. Kembhavi, and A. Schwing. (2021). "Bridging the Imitation Gap by Adaptive Insubordination". In: *Advances in Neural Information Processing Systems*. Ed. by A. Beygelzimer, Y. Dauphin, P. Liang, and J. W. Vaughan. URL: https://openreview.net/forum?id=Wlx0DqiUTD_.

Wen, L., S. Zhang, H. E. Tseng, B. Singh, D. Filev, and H. Peng. (2022). "Improved Robustness and Safety for Pre-Adaptation of Meta Reinforcement Learning with Prior Regularization". In: *2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE. 8987–8994.

Wierstra, D., T. Schaul, T. Glasmachers, Y. Sun, J. Peters, and J. Schmidhuber. (2014). "Natural evolution strategies". *The Journal of Machine Learning Research*. 15(1): 949–980.

Williams, R. J. (1992). "Simple statistical gradient-following algorithms for connectionist reinforcement learning". *Machine learning*. 8(3): 229–256.

Woo, H., G. Yoo, and M. Yoo. (2022). "Structure Learning-Based Task Decomposition for Reinforcement Learning in Non-stationary Environments". In: *Proceedings of the AAAI Conference on Artificial Intelligence*. Vol. 36. No. 8. 8657–8665.

Wu, B., F. Xu, Z. He, A. Gupta, and P. K. Allen. (2020). "SQUIRL: Robust and Efficient Learning from Video Demonstration of Long-Horizon Robotic Manipulation Tasks". In: *IROS*.

Wu, Y., X. Chen, C. Wang, Y. Zhang, and K. W. Ross. (2022). "Aggressive Q-Learning with Ensembles: Achieving Both High Sample Efficiency and High Asymptotic Performance". In: *Deep Reinforcement Learning Workshop NeurIPS 2022*. URL: https://openreview.net/forum?id=L9MhPPvPFS.

Wu, Y., M. Ren, R. Liao, and R. Grosse. (2018). "Understanding Short-Horizon Bias in Stochastic Meta-Optimization". In: *International Conference on Learning Representations*. URL: https://openreview.net/forum?id=H1MczcgR-.

Xian, Z., S. Lal, H.-Y. Tung, E. A. Platanios, and K. Fragkiadaki. (2021). "HyperDynamics: Meta-Learning Object and Agent Dynamics with Hypernetworks". In: *International Conference on Learning Representations*. URL: https://openreview.net/forum?id=pHXfe1cOmA.

Xie, A., A. Singh, S. Levine, and C. Finn. (2018). “Few-shot goal inference for visuomotor learning and planning”. In: *Conference on Robot Learning*. PMLR. 40–52.

Xiong, Z., L. M. Zintgraf, J. A. Beck, R. Vuorio, and S. Whiteson. (2021). “On the Practical Consistency of Meta-Reinforcement Learning Algorithms”. In: *Fifth Workshop on Meta-Learning at the Conference on Neural Information Processing Systems*. URL: https://openreview.net/forum?id=xwQgKphwhFA.

Xu, K., E. Ratner, A. Dragan, S. Levine, and C. Finn. (2019). “Learning a Prior over Intent via Meta-Inverse Reinforcement Learning”. In: *Proceedings of the 36th International Conference on Machine Learning*.

Xu, M., Y. Shen, S. Zhang, Y. Lu, D. Zhao, J. Tenenbaum, and C. Gan. (2022). “Prompting decision transformer for few-shot policy generalization”. In: *international conference on machine learning*. PMLR. 24631–24645.

Xu, Z., H. P. van Hasselt, M. Hessel, J. Oh, S. Singh, and D. Silver. (2020). “Meta-Gradient Reinforcement Learning with an Objective Discovered Online”. In: *Advances in Neural Information Processing Systems*. Ed. by H. Larochelle, M. Ranzato, R. Hadsell, M. F. Balcan, and H. Lin. Vol. 33. Curran Associates, Inc. 15254–15264. URL: https://proceedings.neurips.cc/paper/2020/file/ae3d525daf92cee0003a7f2d92c34ea3-Paper.pdf.

Xu, Z., H. P. van Hasselt, and D. Silver. (2018). “Meta-Gradient Reinforcement Learning”. In: *Advances in Neural Information Processing Systems*. Ed. by S. Bengio, H. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi, and R. Garnett. Vol. 31. Curran Associates, Inc. URL: https://proceedings.neurips.cc/paper/2018/file/2715518c875999308842e3455eda2fe3-Paper.pdf.

Yan, L., D. Liu, Y. Song, and C. Yu. (2020). “Multimodal Aggregation Approach for Memory Vision-Voice Indoor Navigation with Meta-Learning”. In: *IROS*.

Yang, J., E. Wang, R. Trivedi, T. Zhao, and H. Zha. (2022). “Adaptive Incentive Design with Multi-Agent Meta-Gradient Reinforcement Learning”. *AAMAS*.

Yang, J., W. Hu, J. D. Lee, and S. S. Du. (2020). "Impact of representation learning in linear bandits". *arXiv preprint arXiv:2010.06531*.

Yang, Y., K. Caluwaerts, A. Iscen, J. Tan, and C. Finn. (2019). "NoRML: No-Reward Meta Learning". *AAMAS*.

Ye, H., X. Chen, L. Wang, and S. S. Du. (2023). "On the power of pre-training for generalization in rl: Provable benefits and hardness". In: *International Conference on Machine Learning*. PMLR. 39770–39800.

Yin, H., J. Chen, S. J. Pan, and S. Tschiatschek. (2021). "Sequential generative exploration model for partially observable reinforcement learning". In: *Proceedings of the AAAI Conference on Artificial Intelligence*. Vol. 35. No. 12. 10700–10708.

Yoon, J., T. Kim, O. Dia, S. Kim, Y. Bengio, and S. Ahn. (2018). "Bayesian Model-Agnostic Meta-Learning". In: *Advances in Neural Information Processing Systems*. Ed. by S. Bengio, H. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi, and R. Garnett. Vol. 31. Curran Associates, Inc. URL: https://proceedings.neurips.cc/paper/2018/file/e1021d43911ca2c1845910d84f40aeae-Paper.pdf.

Yu, L., T. Yu, C. Finn, and S. Ermon. (2019). "Meta-Inverse Reinforcement Learning with Probabilistic Context Variables". In: *Advances in Neural Information Processing Systems*. Ed. by H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. Fox, and R. Garnett. Vol. 32. Curran Associates, Inc. URL: https://proceedings.neurips.cc/paper/2019/file/30de24287a6d8f07b37c716ad51623a7-Paper.pdf.

Yu, T., C. Finn, S. Dasari, A. Xie, T. Zhang, P. Abbeel, and S. Levine. (2018). "One-Shot Imitation from Observing Humans via Domain-Adaptive Meta-Learning". In: *Proceedings of Robotics: Science and Systems*. Pittsburgh, Pennsylvania. DOI: 10.15607/RSS.2018.XIV.002.

Yu, T., D. Quillen, Z. He, R. Julian, K. Hausman, C. Finn, and S. Levine. (2020a). "Meta-world: A benchmark and evaluation for multi-task and meta reinforcement learning". In: *Conference on Robot Learning*. PMLR. 1094–1100.

Yu, W., J. Tan, Y. Bai, E. Coumans, and S. Ha. (2020b). "Learning fast adaptation with meta strategy optimization". *IEEE Robotics and Automation Letters*. 5(2): 2950–2957.

Yu, W., J. Tan, C. K. Liu, and G. Turk. (2017). "Preparing for the Unknown: Learning a Universal Policy with Online System Identification". In: *Proceedings of Robotics: Science and Systems*. Cambridge, Massachusetts. DOI: 10.15607/RSS.2017.XIII.048.

Yuan, H. and Z. Lu. (2022). "Robust task representations for offline meta-reinforcement learning via contrastive learning". In: *International Conference on Machine Learning*. PMLR. 25747–25759.

Zahavy, T., Z. Xu, V. Veeriah, M. Hessel, J. Oh, H. van Hasselt, D. Silver, and S. Singh. (2020). "A self-tuning actor-critic algorithm". *arXiv preprint arXiv:2002.12928*.

Zang, X., H. Yao, G. Zheng, N. Xu, K. Xu, and Z. Li. (2020). "Metalight: Value-based meta-reinforcement learning for traffic signal control". In: *Proceedings of the AAAI Conference on Artificial Intelligence*. Vol. 34. No. 01. 1153–1160.

Zhang, H. and Z. Kan. (2022). "Temporal logic guided meta q-learning of multiple tasks". *IEEE Robotics and Automation Letters*. 7(3): 8194–8201.

Zhang, J., J. Wang, H. Hu, T. Chen, Y. Chen, C. Fan, and C. Zhang. (2021). "MetaCURE: Meta Reinforcement Learning with Empowerment-Driven Exploration". In: *Proceedings of the 38th International Conference on Machine Learning*. Ed. by M. Meila and T. Zhang. Vol. 139. *Proceedings of Machine Learning Research*. PMLR. 12600–12610. URL: https://proceedings.mlr.press/v139/zhang21w.html.

Zhang, Q. and D. Chen. (2021). "A Meta-Gradient Approach to Learning Cooperative Multi-Agent Communication Topology". In: *Fifth Workshop on Meta-Learning at the Conference on Neural Information Processing Systems*.

Zhao, M., P. Abbeel, and S. James. (2022a). "On the Effectiveness of Fine-tuning Versus Meta-reinforcement Learning". In: *Advances in Neural Information Processing Systems*. Ed. by A. H. Oh, A. Agarwal, D. Belgrave, and K. Cho. URL: https://openreview.net/forum?id=mux7gn3g_3.

Zhao, T. Z., J. Luo, O. Sushkov, R. Pevceviciute, N. Heess, J. Scholz, S. Schaal, and S. Levine. (2022b). "Offline meta-reinforcement learning for industrial insertion". In: *2022 International Conference on Robotics and Automation (ICRA)*. IEEE. 6386–6393.

Zhao, W., J. P. Queralta, and T. Westerlund. (2020). "Sim-to-real transfer in deep reinforcement learning for robotics: a survey". In: *2020 IEEE Symposium Series on Computational Intelligence (SSCI)*. IEEE. 737–744.

Zhao, Z., A. Nagabandi, K. Rakelly, C. Finn, and S. Levine. (2021). "MELD: Meta-Reinforcement Learning from Images via Latent State Models". In: *Conference on Robot Learning*. PMLR. 1246–1261.

Zheng, Z., J. Oh, M. Hessel, Z. Xu, M. Kroiss, H. Van Hasselt, D. Silver, and S. Singh. (2020). "What can learned intrinsic rewards capture?" In: *International Conference on Machine Learning*. PMLR. 11436–11446.

Zheng, Z., J. Oh, and S. Singh. (2018). "On Learning Intrinsic Rewards for Policy Gradient Methods". In: *Advances in Neural Information Processing Systems*. Ed. by S. Bengio, H. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi, and R. Garnett. Vol. 31. Curran Associates, Inc. URL: https://proceedings.neurips.cc/paper/2018/file/51de85ddd068f0bc787691d356176df9-Paper.pdf.

Zhou, A., E. Jang, D. Kappler, A. Herzog, M. Khansari, P. Wohlhart, Y. Bai, M. Kalakrishnan, S. Levine, and C. Finn. (2020). "Watch, Try, Learn: Meta-Learning from Demonstrations and Rewards". In: *ICLR*.

Zhou, W., L. Pinto, and A. Gupta. (2019). "Environment Probing Interaction Policies". In: *International Conference on Learning Representations*. URL: https://openreview.net/forum?id=ryl8-3AcFX.

Zintgraf, L., S. Devlin, K. Ciosek, S. Whiteson, and K. Hofmann. (2021a). "Deep Interactive Bayesian Reinforcement Learning via Meta-Learning". In: *Proceedings of the 20th International Conference on Autonomous Agents and MultiAgent Systems*. 1712–1714.

Zintgraf, L., S. Schulze, C. Lu, L. Feng, M. Igl, K. Shiarlis, Y. Gal, K. Hofmann, and S. Whiteson. (2021b). "VariBAD: Variational Bayes-Adaptive Deep RL via Meta-Learning". *Journal of Machine Learning Research*. 22(289): 1–39.

Zintgraf, L., K. Shiarlis, M. Igl, S. Schulze, Y. Gal, K. Hofmann, and S. Whiteson. (2020). “VariBAD: A Very Good Method for Bayes-Adaptive Deep RL via Meta-Learning”. In: *International Conference on Learning Representation (ICLR)*.

Zintgraf, L. M., L. Feng, C. Lu, M. Igl, K. Hartikainen, K. Hofmann, and S. Whiteson. (2021c). “Exploration in Approximate Hyper-State Space for Meta Reinforcement Learning”. In: *International Conference on Machine Learning*. PMLR. 12991–13001.

Zintgraf, L. M., K. Shiarlis, V. Kurin, K. Hofmann, and S. Whiteson. (2019). “Fast Context Adaptation via Meta-Learning”. *ICLR*.

Zou, H., T. Ren, D. Yan, H. Su, and J. Zhu. (2021). “Learning task-distribution reward shaping with meta-learning”. In: *Proceedings of the AAAI Conference on Artificial Intelligence*. Vol. 35. No. 12. 11210–11218.

Zou, Y. and X. Lu. (2020). “Gradient-EM Bayesian Meta-Learning”. In: *Advances in Neural Information Processing Systems*. Ed. by H. Larochelle, M. Ranzato, R. Hadsell, M. Balcan, and H. Lin. Vol. 33. Curran Associates, Inc. 20865–20875. URL: https://proceedings.neurips.cc/paper/2020/file/ef48e3ef07e359006f7869b04fa07f5e-Paper.pdf.