

OTTO: A Python package to simulate, solve and visualize the source-tracking POMDP

Aurore Loisy^{1¶} and Christophe Eloy^{1¶}

¹ Aix Marseille Univ, CNRS, Centrale Marseille, IRPHE, Marseille, France ¶ Corresponding author

DOI: [10.xxxxxx/draft](https://doi.org/10.xxxxxx/draft)

Software

- [Review](#) ↗
- [Repository](#) ↗
- [Archive](#) ↗

Editor: ↗

Reviewers:

- [@11michalis11](#)

Submitted: 07 March 2022

Published: unpublished

License

Authors of papers retain copyright and release the work under a Creative Commons Attribution 4.0 International License ([CC BY 4.0](#)).

Statement of need

The source-tracking problem is a POMDP (partially observable Markov decision process) designed by Vergassola et al. (2007) to mimic the problem of searching for a source of odor in a turbulent flow. Far from a “toy” POMDP, it incorporates physical models of odor dispersion and detection that reproduce the major features of olfactory searches in turbulence. Solutions to this problem, even approximate, have direct applications to sniffer robots used to track chemicals emitted by explosives, drugs or chemical leaks (Marques & Almeida, 2006; Russell, 1999). They may also shed light on how cognitive animals use olfaction to search for food and mates (Reddy et al., 2022; Vickers, 2000).

In the source-tracking POMDP, the agent must find a source of odor hidden in a grid world. At each step, the agent moves to a neighbor cell. Once in a new cell, the agent receives an observation (odor detection) that provides some noisy information on its current distance to the source. The search terminates when the agent enters the cell containing the source. Solving the POMDP means finding the optimal way of choosing moves (policy) so as to reach the source in the smallest possible number of steps.

Computing the optimal policy is not possible, and the challenge resides in finding good approximate solutions. A strong baseline is provided by “infotaxis”, a heuristic policy devised by Vergassola et al. (2007). It has become popular in robotics (Lochmatter, 2010; Moraud & Martinez, 2010) and to interpret animal searches (Calhoun et al., 2014; Vergassola et al., 2007; Voges et al., 2014).

Several variants have been proposed since (Chen et al., 2020; Hutchinson et al., 2018; Karpas et al., 2017; Masson, 2013; Ristic et al., 2016), but the quest for better policies has been hindered by the lack of a trustable, open-source implementation of the source-tracking POMDP. Existing comparisons suffer from this lack of a common implementation. Besides, the recent successes of reinforcement learning for solving complex navigation problems in turbulence (Alageshan et al., 2020; Reddy et al., 2016) calls for an adaptation of these methods to the source-tracking POMDP.

The target audience of OTTO consists of researchers in biophysics, applied mathematics and robotics working on optimal strategies for olfactory searches in turbulent flows.

Summary

OTTO (short for Odor-based Target Tracking Optimization) is a Python package to learn, evaluate and visualize strategies for odor-based searches.

OTTO provides:

1. a simulator of the source-tracking POMDP for any number of space dimensions, domain sizes and source intensities, together with a rendering tool in 1D, 2D and 3D;

2. an implementation of several heuristic policies including “infotaxis” (Vergassola et al., 2007) and its recently proposed “space-aware” variant (Loisy & Eloy, 2021);
3. a parallelized algorithm to evaluate policies (probability of finding the source, distribution of search times, etc.) using a rigorous, well-defined protocol;
4. a custom model-based deep reinforcement learning algorithm for training neural-network policies, together with a library (“zoo”) of trained neural networks that achieve near-optimal performance;
5. a wrapper of the source-tracking POMDP that follows the OpenAI Gym interface.

OTTO aims at facilitating future research:

1. New heuristic policies can easily be implemented, visualized, and evaluated. To facilitate comparison to existing baselines, the performance of several policies (including infotaxis and near-optimal) is reported in a freely available dataset generated with OTTO (Loisy & Eloy, 2022).
2. The gym wrapper makes the source-tracking POMDP easily accessible to the reinforcement learning community. OpenAI Gym (Brockman et al., 2016) is the *de facto* standard for simulators. It is compatible with most general-purpose model-free reinforcement learning libraries (e.g., Stable Baselines (Raffin et al., 2021), OpenAI-Baselines (Dhariwal et al., 2017), RLlib (Liang et al., 2018), CleanRL (Huang et al., 2021), ChainerRL/PFRL (Fujita et al., 2021)).

Mentions

The methodological aspects of OTTO (generalization of the POMDP to an arbitrary number of space dimensions, policy evaluation protocol, model-based reinforcement learning algorithm) have been developed as part of a publication by its authors (Loisy & Eloy, 2021).

Acknowledgements

This project has received funding from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme (grant agreement No 834238).

References

- Alageshan, J. K., Verma, A. K., Bec, J., & Pandit, R. (2020). Machine learning strategies for path-planning microswimmers in turbulent flows. *Physical Review E*, 101(4), 043110. <https://doi.org/10.1103/PhysRevE.101.043110>
- Brockman, G., Cheung, V., Pettersson, L., Schneider, J., Schulman, J., Tang, J., & Zaremba, W. (2016). *OpenAI gym*. <https://arxiv.org/abs/1606.01540>
- Calhoun, A. J., Chalasani, S. H., & Sharpee, T. O. (2014). Maximally informative foraging by *Caenorhabditis elegans*. *eLife*, 3, e04220. <https://doi.org/10.7554/eLife.04220>
- Chen, C., Murphey, T. D., & MacIver, M. A. (2020). Tuning movement for sensing in an uncertain world. *eLife*, 9, e52371. <https://doi.org/10.7554/elife.52371>
- Dhariwal, P., Hesse, C., Klimov, O., Nichol, A., Plappert, M., Radford, A., Schulman, J., Sidor, S., Wu, Y., & Zhokhov, P. (2017). OpenAI baselines. In *GitHub repository*. <https://github.com/openai/baselines>; GitHub.
- Fujita, Y., Nagarajan, P., Kataoka, T., & Ishikawa, T. (2021). ChainerRL: A deep reinforcement learning library. *Journal of Machine Learning Research*, 22(77), 1–14. <http://jmlr.org/papers/v22/20-376.html>

- 83 Huang, S., Dossa, R. F. J., Ye, C., & Braga, J. (2021). *CleanRL: High-quality single-file*
84 *implementations of deep reinforcement learning algorithms*. [https://arxiv.org/abs/2111.](https://arxiv.org/abs/2111.08819)
85 [08819](https://arxiv.org/abs/2111.08819)
- 86 Hutchinson, M., Oh, H., & Chen, W.-H. (2018). Entrotaxis as a strategy for autonomous
87 search and source reconstruction in turbulent conditions. *Information Fusion*, 42, 179–189.
88 <https://doi.org/10.1016/J.INFFUS.2017.10.009>
- 89 Karpas, E. D., Shklarsh, A., & Schneidman, E. (2017). Information socialtaxis and efficient
90 collective behavior emerging in groups of information-seeking agents. *Proceedings of*
91 *the National Academy of Sciences*, 114(22), 5589–5594. [https://doi.org/10.1073/pnas.](https://doi.org/10.1073/pnas.1618055114)
92 [1618055114](https://doi.org/10.1073/pnas.1618055114)
- 93 Liang, E., Liaw, R., Nishihara, R., Moritz, P., Fox, R., Goldberg, K., Gonzalez, J., Jordan,
94 M., & Stoica, I. (2018). RLlib: Abstractions for distributed reinforcement learning.
95 *Proceedings of the 35th International Conference on Machine Learning*, 80, 3053–3062.
96 <https://proceedings.mlr.press/v80/liang18b.html>
- 97 Lochmatter, T. (2010). *Bio-inspired and probabilistic algorithms for distributed odor source*
98 *localization using mobile robots* [PhD thesis]. EPFL.
- 99 Loisy, A., & Eloy, C. (2021). *Searching for a source without gradients: How good is infotaxis*
100 *and how to beat it*. <https://arxiv.org/abs/2112.10861>
- 101 Loisy, A., & Eloy, C. (2022). Dataset for "Searching for a source without gradients: how good
102 is infotaxis and how to beat it". <https://doi.org/10.5281/zenodo.6125391>
- 103 Marques, L., & Almeida, A. de. (2006). Mobile robot olfaction. *Autonomous Robots*, 20,
104 183–184. <https://doi.org/10.1007/s10514-006-7536-7>
- 105 Masson, J.-B. (2013). Olfactory searches with limited space perception. *Proceedings of the*
106 *National Academy of Sciences of the United States of America*, 110(28), 11261–11266.
107 <https://doi.org/10.1073/pnas.1221091110>
- 108 Moraud, E. M., & Martinez, D. (2010). Effectiveness and robustness of robot infotaxis for
109 searching in dilute conditions. *Frontiers in Neurorobotics*, 4, 1–8. [https://doi.org/10.3389/](https://doi.org/10.3389/fnbot.2010.00001)
110 [fnbot.2010.00001](https://doi.org/10.3389/fnbot.2010.00001)
- 111 Raffin, A., Hill, A., Gleave, A., Kanervisto, A., Ernestus, M., & Dormann, N. (2021). Stable-
112 Baselines3: Reliable reinforcement learning implementations. *Journal of Machine Learning*
113 *Research*, 22(268), 1–8. <http://jmlr.org/papers/v22/20-1364.html>
- 114 Reddy, G., Celani, A., Sejnowski, T. J., & Vergassola, M. (2016). Learning to soar in turbulent
115 environments. *Proceedings of the National Academy of Sciences of the United States of*
116 *America*, 113(33), E4877–84. <https://doi.org/10.1073/pnas.1606075113>
- 117 Reddy, G., Murthy, V. N., & Vergassola, M. (2022). Olfactory Sensing and Navigation
118 in Turbulent Environments. *Annual Review of Condensed Matter Physics*, 13(1), tbd.
119 <https://doi.org/10.1146/annurev-conmatphys-031720-032754>
- 120 Ristic, B., Skvortsov, A., & Gunatilaka, A. (2016). A study of cognitive strategies for an
121 autonomous search. *Information Fusion*, 28, 1–9. [https://doi.org/10.1016/J.INFFUS.2015.](https://doi.org/10.1016/J.INFFUS.2015.06.008)
122 [06.008](https://doi.org/10.1016/J.INFFUS.2015.06.008)
- 123 Russell, R. A. (1999). *Odour detection by mobile robots*. World Scientific. [https://doi.org/](https://doi.org/10.1142/4042)
124 [10.1142/4042](https://doi.org/10.1142/4042)
- 125 Vergassola, M., Villermaux, E., & Shraiman, B. I. (2007). "Infotaxis" as a strategy for searching
126 without gradients. *Nature*, 445(7126), 406–409. <https://doi.org/10.1038/nature05464>
- 127 Vickers, N. J. (2000). Mechanisms of animal navigation in odor plumes. *The Biological*
128 *Bulletin*, 198(2), 203–212. <https://doi.org/10.2307/1542524>

- ¹²⁹ Voges, N., Chaffiol, A., Lucas, P., & Martinez, D. (2014). Reactive Searching and Infotaxis
¹³⁰ in Odor Source Localization. *PLOS Computational Biology*, 10(10), e1003861. [https:](https://doi.org/10.1371/journal.pcbi.1003861)
¹³¹ [//doi.org/10.1371/journal.pcbi.1003861](https://doi.org/10.1371/journal.pcbi.1003861)

DRAFT