Data-driven Bayesian olfactory search in a turbulent flow

R. A. Heinonen[1], F. Bonaccorso[1], L. Biferale[1], A. Celani[2], and M. Vergassola[3]

[1]Dept. Physics and INFN, University of Rome, “Tor Vergata”
[2]The Abdus Salam International Center for Theoretical Physics
[3]Dept. Physics, Ecole Normale Supérieure

Supported by the European Research Council under grant No. 882340

Introduction: searching for an odor source in a turbulent environment

- Insects often need find source (usually upwind) of an odor or other cue advected by the atmosphere
- E.g. mosquito looking for human drawn by CO_2 and odors; moth looking for mate drawn by pheromones
- Source may be $\sim$ 100 m away(!)

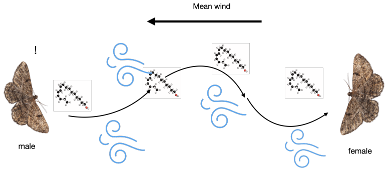


Figure Artist's conception of a moth searching for a mate via pheromone cues.

Introduction: searching for an odor source in a turbulent environment

- Classical search strategy is chemotaxis, i.e. just go up the concentration gradient
- But: (far from source) turbulence mixes cue into patches/plumes over background of very small concentration $\implies$ insect only detects the cue **intermittently**. Gradient estimation is unfeasible

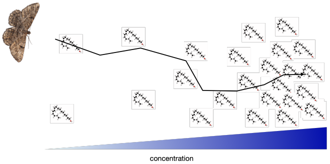


Figure Artist's conception of chemotaxis strategy.

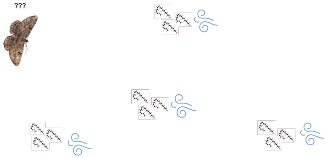


Figure A turbulent environment leads to a patchy odor landscape with intermittent detections.

Model search problem

- Agent makes observation — detection or nondetection, then moves
- Try to reach source in as few Δt as possible — give reward γ^T for reaching source in T steps ($0 < \gamma < 1$)
- Key physics input is $\mathrm{Pr}(\mathrm{obs}|\mathbf{s}), \mathbf{r} - \mathbf{r}_0$. Spatial dependence of concentration statistics in turbulent environment? (see [Celani et al., 2014])

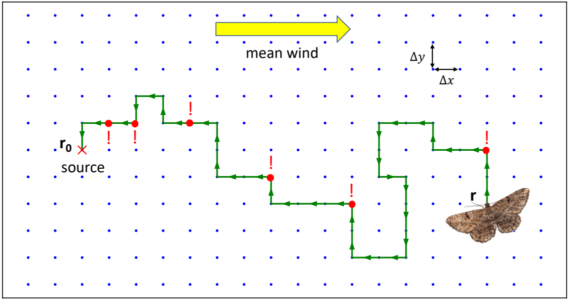


Figure In our setup, agent lives on the gridworld (blue points) and tries to find the source (red x)

Capturing the information

- At timestep t, agent has history $(a_1, o_1, a_2, o_2, \ldots, a_{t-1}, o_t)$. What does this say about source location?
- Assuming system is Markovian, information can be stored in a probability distribution ("belief") b over $\mathbf{s}$
- Update b after each observation using Bayes' theorem

$$b(s')_{o,a} = \mathrm{Pr}(o|s') \sum_s b(s)\mathrm{Pr}(s'|s,a)/Z$$

- This describes a partially observable Markov decision process (POMDP) — state not accessible to agent, only observations
- *Model-based* approach — need $\mathrm{Pr}(o|\mathbf{s})$

Optimal policy: Bellman equation

- Define value function $V_\pi(b)$ as total expected reward $\mathbb{E}[\gamma^T]$ under π, conditioned on b. Optimal value function satisfies Bellman equation

$$V^*(b) = \max_{a \in A} \left[\underbrace{\sum_{s \in S} R(s,a) b(s)}_{\substack{\text{immediate expected} \\ \text{reward}}} + \gamma \underbrace{\sum_{o \in O} \Pr(o|b,a) V^*(b_{o,a})}_{\text{future expected rewards}} \right]$$

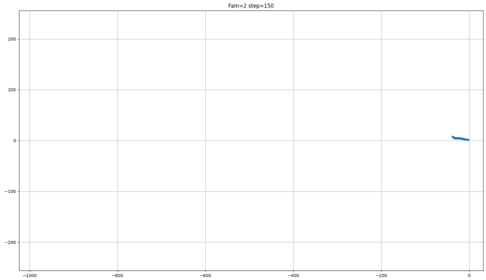
- Partial observability makes solution computationally hard — belief simplex very large (dimension $|S| - 1$). "Curse of dimensionality"

Previous work

- Recent work has shown this problem can be solved effectively using at least three algorithms (Perseus w/ reward shaping, SARSOP, stochastic gradient descent)
    1. Loisy and Eloy *Proc. R. Soc. Lond.* (2022)
    2. Heinonen, Biferale et al. (2022, under review)
    3. Loisy, Heinonen et al. (in preparation)
- This research focused on the “toy problem” where the model is exact
- Now we move to a “real” turbulent flow (DNS)!

The DNS

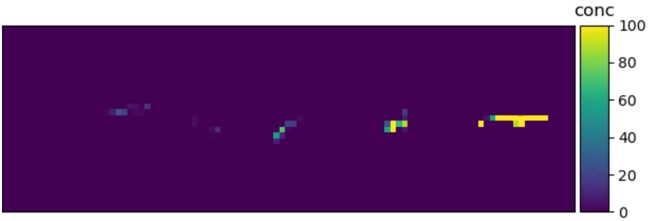
- 3-D Navier-Stokes with mean wind on $1024 \times 512 \times 512$ grid, $\mathrm{Re} \simeq 750$
- Lagrangian particles emitted simultaneously from 5 different sources, data outputted every τ_η ($\sim 5000\tau_\eta$ total)
- Have data for 5 different wind speeds ($V/\tilde{v} \simeq 0, 1.5, 3, 6, 9$)
- Let us know if you have interest in this dataset!



N.B. our simulation guy got the flu and didn't finish the 3-D movie

Coarse-graining

- To move to POMDP setting, data are coarse-grained on a quasi-2D slice to obtain 99×33 grid, spacing is $\sim 10\eta$
- Particles counted to obtain concentration field



N.B. these movies have $V = 6$, data in the rest of talk are for $V = 9$

Thresholding

- Agent moves 1 square per τ_η, observes instantaneous concentration c
- Hit defined as $c \geq c_{thr} = 100$

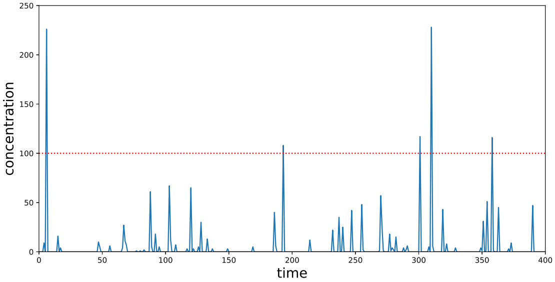


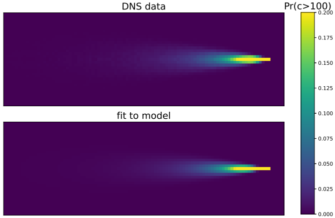
Figure Conc. time series at fixed point $58\Delta x$ downwind from the source, with detection threshold

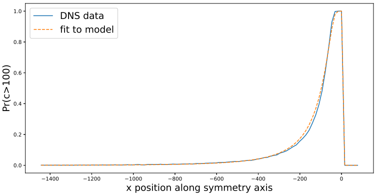
Empirical likelihood

- $\Pr(c \geq c_{thr} | \mathbf{s})$ averaged over time and source locations, symmetrized across wind axis
- Now, fit to two-parameter model based on [Celani et al., 2014]:

$$\Pr(c \geq c_{thr} | x, y) =$$
$$\theta(x)\left(1 - \exp\left(\chi \, \mathrm{Ei}(-b/x)\right)\right),$$
$$\chi = \frac{a}{x} \exp\left[-\left(\frac{V}{\tilde{v}_\perp}\right)^2 \left(\frac{y}{x}\right)^2\right]$$

- Use SARSOP (or other) to solve for policy assuming fit model is exact

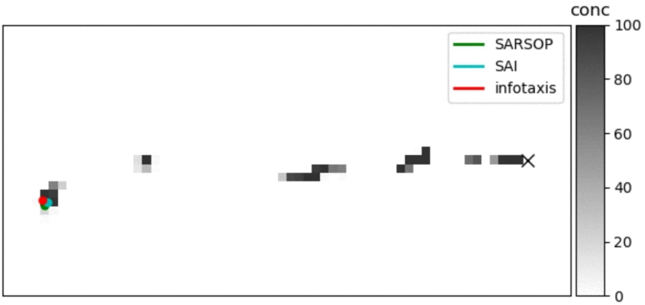




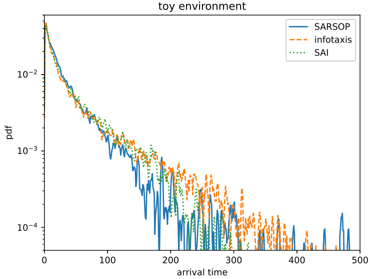
N.B. $\mathrm{Ei}(x) = \int_{-\infty}^{x} \frac{e^t}{t} \, dt$

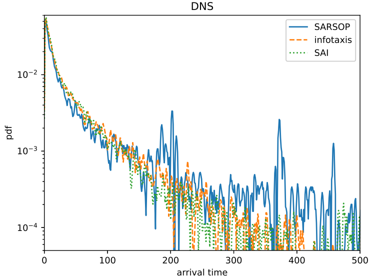
Searching in the DNS: near-optimal vs. heuristics

Near-optimal policy (SARSOP) generally outperforms the two tested heuristic (i.e. not optimized) policies when starting reasonably far from the source



Arrival time statistics





policy	toy env.	DNS
SARSOP	36.9	125.6
SAI	40.2	38.9
infotaxis	47.8	41.5

Table Mean arrival times $\mathbb{E}[T \,|\, T < 2500]$

policy	toy env.	DNS
SARSOP	0%	0.55%
SAI	0.37%	5.8%
infotaxis	0.22%	5.2%

Table Failure rates $\Pr(T \geq 2500)$

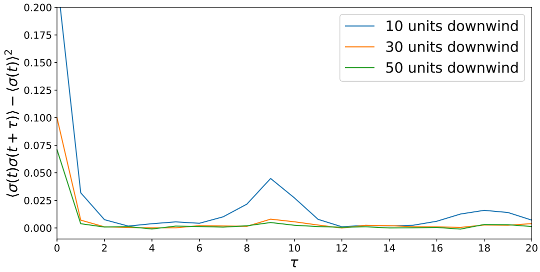
Failure rate larger and tails fatter in DNS

Correlations between detections

- Simplest explanation for discrepancy: real turbulence is not Markovian. Detections are positively correlated in time in DNS
- Consecutive detections more likely than in Markov model $\Longrightarrow$ POMDP agent sometimes "fooled"
- Define binary signal $\sigma(t) = \theta(c(t) - c_{thr})$, then

$$\langle \sigma(t+\tau)\sigma(t) \rangle - \langle \sigma(t) \rangle^2 = \chi \mathrm{Pr}(\text{det. at } t+\tau | \text{det. at } t) - \chi^2$$

where $\chi = \mathrm{Pr}(c \geq c_{thr})$. Expect $\propto \delta_{\tau,0}$ if uncorrelated



Nonzero positive correlations, especially close to the source! (Here $c_{thr} = 10$ for better statistics)

Conclusions

- Have high-quality Lagrangian data for particles in a turbulent flow emitted from point source in the presence of mean wind
- Have used the data + POMDP algorithms to solve for near-optimal policies for olfactory search
- Correlations between detections can spoil performance
- Next steps:
    1. How does a model-free policy with memory compare to POMDP?
    2. Move to *risk-averse* setting: try to minimize $\mathbb{E}[\exp(\beta T)]$ for $\beta > 0 \implies$ optimize the tail and avoid failure
    3. Performance sensitive to model parameters, but model not always available in real world. How to relax this?

References I

Celani, A., Villermaux, E., and Vergassola, M. (2014).
Odor landscapes in turbulent environments.
Physical Review X, 4(4):041015.

Kaelbling, L. P., Littman, M. L., and Cassandra, A. R. (1998).
Planning and acting in partially observable stochastic domains.
Artificial intelligence, 101(1-2):99–134.

Loisy, A. and Eloy, C. (2021).
Searching for a source without gradients: how good is infotaxis and how to beat it.
arXiv preprint arXiv:2112.10861.

Vergassola, M., Villermaux, E., and Shraiman, B. I. (2007).
'Infotaxis' as a strategy for searching without gradients.
Nature, 445(7126):406–409.

Detection likelihood model details

- [Celani et al., 2014] calculated concentration statistics far from a source in the presence of wind
- Concentration of a puff controlled by the size of the puff when it passed through the source
- Compute puff size statistics, prob. of passing through source $\rightarrow$ can compute $\chi = \mathrm{Pr}(c > 0)$ and $C = \langle c | c > 0 \rangle$
- For jet flow and Gaussian fluctuations, obtain $\chi \propto \exp(-(Vy/\tilde{v}x)^2)/x$, $C \propto 1/x$
- Poisson statistics for rare events $\rightarrow$ tail of pdf shown to be $p(c) \sim \frac{\chi}{c} \exp(-c/C)$
- Can integrate to find $\mathrm{Pr}(c \geq c_{thr})$

Heuristic strategies

- QMDP: take action which essentially minimizes the expected distance to the source. Exploitative (greedy)
- Infotaxis [Vergassola et al., 2007]: take action maximizing the expected gain in information (negative entropy) $I = \sum_s b(s) \log b(s)$. Explorative (less greedy)
- Space-aware infotaxis [Loisy and Eloy, 2021]: take action minimizing a function with contributions from both the distance and the entropy
- Thompson sampling: sample a point $\mathbf{r}^*$ from b, move for τ timesteps towards $\mathbf{r}^*$, repeat.