Learning Latent Representations for Active Search on Partially Observable Graphs

Author: Stefan Voicila • Supervisors: Prof. Dr. Frans A. Oliehoek, Andreea Minculescu

TU Delft
Research Project
TU Delft 2026

1. Motivation

- Active search: find target nodes under a query budget.
- Prior work assumes known topology.
- Both unknown → POMDP over graph–label pairs.
- Particle solvers (POMCP) store full graphs, costly as *n* grows.

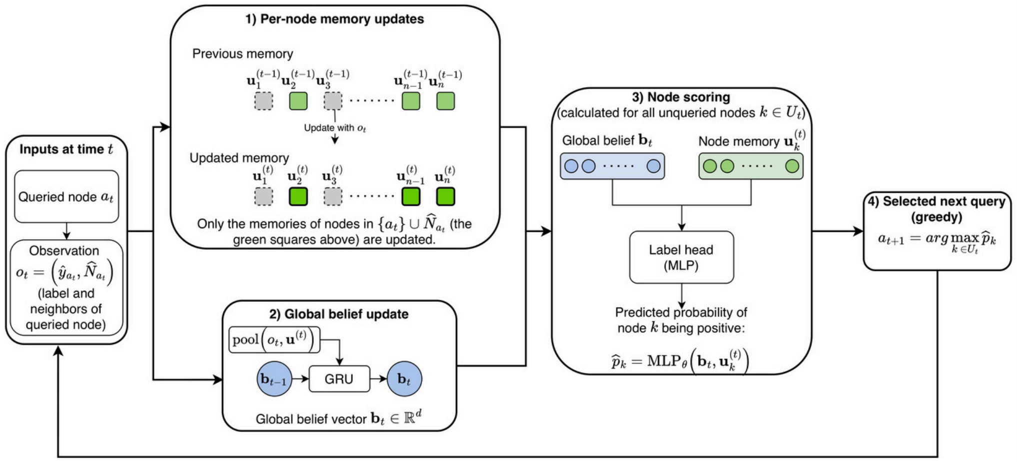
Research question. Can a low-dimensional learned belief over the hidden graph be maintained tractably, while still supporting effective active-search decisions?

2. Key idea & contribution

- Replace the particle belief with a learned latent belief $\mathbf{b}_t$, updated directly from observations.
- A global belief vector $\mathbf{b}_t$ captures graph-level structure, while per-node memories store node-specific evidence.
- Nodes are scored using both representations, and the highest-scoring node is queried.

Contribution. A particle belief stores an entire graph per hypothesis and grows rapidly with graph size. In contrast, the proposed latent belief has a fixed dimension d, while only the per-node memory scales linearly with the number of nodes *n*.

3. Architecture



One inference step: the queried node's observation updates per-node memories (green) and the global belief $\mathbf{b}_t$; the highest-scoring unqueried node is queried next.

4. Results

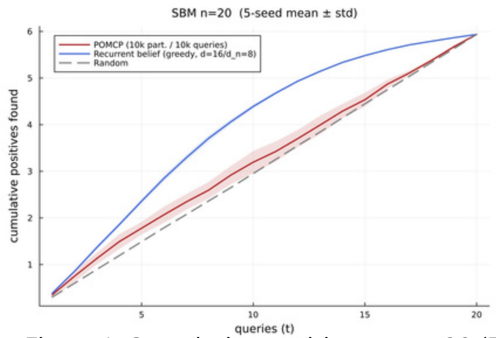


Figure 1. Cumulative positives at *n* = 20 (5-seed mean ± std). The recurrent belief leads throughout the intermediate-budget regime - at 10 queries it finds 4.4 positives, versus 3.2 (POMCP) and 3.0 (random).

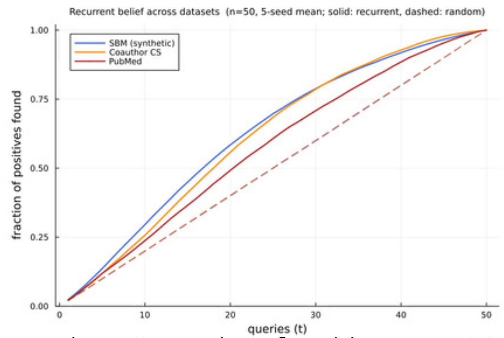


Figure 2. Fraction of positives at *n* = 50 on synthetic SBM and two real graphs. The learned belief transfers to real data, staying above random throughout.

Half-budget recall ($B = n/2$): 0.74/0.70/0.67 at n = 20/50/75 vs ≈0.50 random. The lead holds as the graph grows, exactly where particle POMCP becomes intractable.

5. Conclusions

- Compact latent belief replaces graph particles - fixed size, no graph reconstruction.
- Beats POMCP at *n*=20, beats random at *n*=50/75, transfers to two real graphs.

Limitations

- Evaluated mostly on synthetic SBM; only two real graphs, both *n*=50.
- Training is supervised rather than reinforcement learning, so the greedy policy is implicitly exploitative and does not explicitly trade off exploration.

Future work

- Cross-distribution transfer (Barabási–Albert, small-world); zero-shot = the strong claim.
- Larger, heavier-tailed real graphs (e.g. fraud-detection networks)
- A reinforcement-learning objective for explicit exploration.
- Calibration analysis of the label-head probabilities under focal loss.

Key References: [1] Garnett, R., et al. Bayesian Optimal Active Search and Surveying. ICML, 2012. [2] Wang, X., et al. Active Search on Graphs. KDD, 2013. [3]Jiang, S., et al. Efficient Nonmyopic Active Search. ICML, 2017. [4] Silver, D., and Veness, J. Monte-Carlo Planning in Large POMDPs. NeurIPS, 2010. [5] Hausknecht, M., and Stone, P. Deep Recurrent Q-Learning for Partially Observable MDPs. AAAI Fall Symposium, 2015. [6] Igl, M., et al. Deep Variational Reinforcement Learning for POMDPs. ICML, 2018. [7]Cho, K., et al. Learning Phrase Representations using RNN Encoder–Decoder for Statistical Machine Translation.EMNLP, 2014.