Map Induction: Compositional spatial submap learning for efficient exploration in novel environments

Summary: Humans efficiently reason about space to navigate and forage in new environments, which may, at least in part, depend on an ability to generalize across tasks and organize observations into patterns that can be re-used [1]. Generalization and transfer learning in spatial domains is evident in mirror-invariant neural scene representations [2], and in the reuse of reference frames and representations across similar environments [3]. Shared reference frames appear in other mammals as well, as, for example, in rodents reusing grid-cell maps for different but perceptually similar environments [4-5]. However, the field lacks conceptual and quantitative models of how the spatial system might discover patterns during spatial exploration, how seen patterns might be compositionally combined to represent complex spaces, or how they might be leveraged to represent and navigate through new spaces through reuse. Here we introduce a computational model of "Map Induction", which involves the compositional formation of proposed maps of complex spaces based on already-seen spaces through program induction in a Hierarchical Bayesian framework. The model thus explicitly reasons about unseen spaces through a distribution of strong spatial priors. We introduce a new behavioral Map Induction Task (MIT), and compare human performance with that of state-of-the-art Partially Observable Monte Carlo panning models as well as our Map Induction framework. We show that our computational framework better predicts human exploration behavior than non-inductive models. Understanding the computational mechanisms that support such map learning can generate hypotheses for circuit-level neural representations and dynamics, advance the study of the human mind, as well as support more efficient exploration algorithms.

Further Details: We propose that humans optimize exploration of new spaces by representing maps as composed of reusable reference frames. More specifically, we propose a map induction hypothesis – that humans use program induction to infer possible maps of unseen spaces, as made up of submaps encountered in the observed areas. To study map induction during exploration, we designed the MIT task in which subjects forage for rewards in novel, partially observable 3-dimensional environments (top down views of sample environments shown in Figure 1d). All environments contain multiple rewards (diamonds), hidden in predictable locations within each unit. The subjects have no advance knowledge of environment structure, and are instructed to collect all diamonds while navigating the environments from a first-person perspective.

We formalize exploration in the MIT task as four computational steps (Figure 1a), the first three of which constitute map induction – approximate inference of a posterior distribution $p(M|D)$ over possible maps, given a history of observations. In the fourth step, the inferred map distribution is used to plan exploration. **(1) The Region Extractor** extracts regions (submaps) Mp from previously explored parts of the map (Figure 1b, second level). **(2) The Map Generator** develops a space of possible map completions M by a probabilistic generative grammar assembling the extracted regions into hypothetical maps (Figure 1b, third level). **(3) Map Inference** computes the posterior distribution $p(M|D)$ using a likelihood function, and a prior on complete maps derived from the probabilities of the production rules in the generative grammar. At a high level, the likelihood function encodes two simple assumptions: (1) humans prefer map completions that are maximally descriptive of unseen portions of the environment; (2) humans prefer map hypotheses that are minimally contradictory with previous observations. **(4) Planner:** we model the exploration problem as a Partially Observed Markov Decision Process (POMDP), implemented by an online Partially Observable Monte Carlo Planner (POMCP) [6] and compare humans to three variants of POMCP (Figure 1c).

Experiment 1 (Figure 1d, left) tests whether humans perform map induction (implemented by MAP-POMCP and D-POMCP) in contrast to naive exploration (Uniform-POMCP). Results (Figure 1e) favour the MAP-POMCP and D-POMCP models, which implement map induction, over the Uniform-POMCP model, which makes no predictions about map structure. Experiment 2 (Figure 1d, right) differentiates between MAP-POMCP and D-POMCP using color cues to indicate the location of rewards within each unit. Results (Figure 1f) indicate that D-POMCP model, which plans over a distribution of possible map completions, was more likely to explain human behaviour compared to MAP-POMCP, which plans exclusively using a single map completion with the highest posterior likelihood under the model.

The map induction hypothesis for explaining human exploration is also supported by hippocampal reuse of place-cell maps in composite environments [7], which are invariant to rotation and scaling [8].

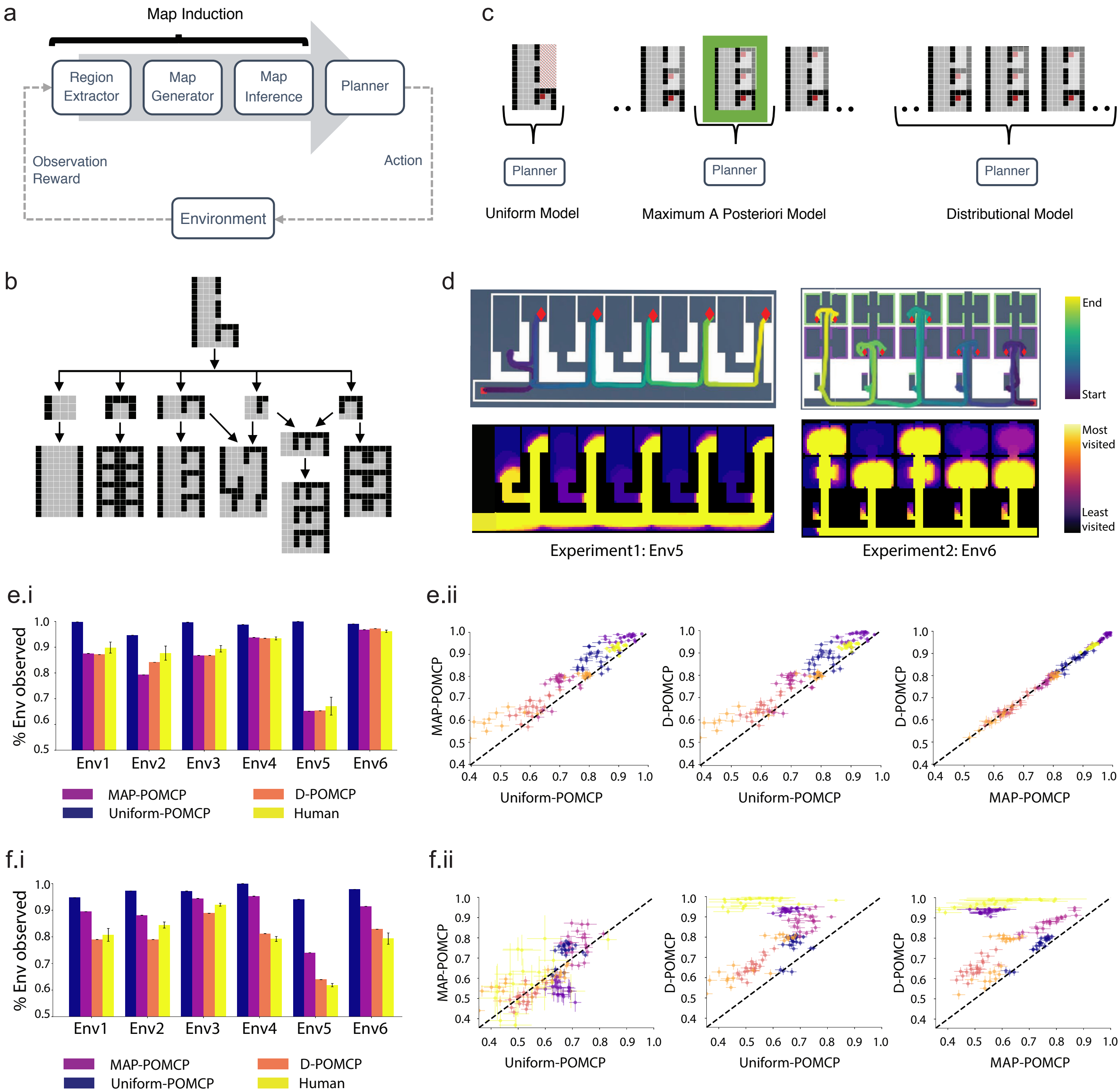

Figure 1: (a) Model Architecture: four steps required to solve the MIT task and the corresponding computational modules in our framework (b) Extraction of candidate regions (submaps) from the observed portion of the environment and generation of possible map completions. (c) Three model hypotheses: Uniform Model assumes a uniform distribution for the unobserved part of the map; Maximum A Posteriori Model uses the most likely map completion for planning; Distributional Model uses the entire posterior distribution for planning. (d) Top: exploration trajectories of a representative subject in Experiment 1 (left) and Experiment 2 (right). Bottom: Visitation heatmaps of an example environment aggregated across subjects in Experiment 1 (left) and Experiment 2 (right). The visited area was computed in a 2D grid projection, using a circular radius of five grid cells around the agent. (e) Experiment 1 results. i) Fractions of environments observed by the models and humans. Error bars show 95% confidence intervals. ii) Percent likelihood that the model takes the same actions as the human. Each marker is a subject-environment pair, with each color showing a single environement: Env1-Env6 (blue-yellow).Error bars show standard error along each axis. (f) Experiment 2 results.

References: [1] Pitt, Benjamin, et al. (2021). [2] Dilks, Daniel D., et al. (2011). [3] Marchette, Steven A., et al. (2014). [4] Derdikman, Dori, et al. (2009). [5] Carpenter, Francis, et al. (2015). [6] Silver, David, and Joel Veness.(2010). [7] Paz-Villagrán, V., E. Save, and Bruno Poucet.(2004). [8] Muller, Robert U., and John L. Kubie. (1987).