Few-Shot Spatial Planning in Humans and Deep Reinforcement Learning Agents

(Zero?)

Denis C. L. Lan, Laurence T. Hunt*, Christopher Summerfield*

Department of Experimental Psychology, University of Oxford

(* Contributed equally as senior authors)

Introduction

- Humans plan flexibly in spatial settings, even with little knowledge and experience (i.e., in a few-shot manner)
- This likely depends on our ability to arbitrate between different planning strategies:

Vector-Based

Dependent on knowledge of relative spatial locations of (landmark) states

Transition-Based

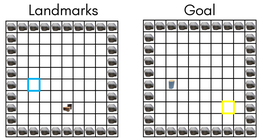
Dependent on learnt knowledge of relationships between states

Hypothesis: We use *vectors* to head in the general goal direction, but *transitions* to fine-tune navigation in familiar regions of state space

Task Design

'Few-shot' navigation task: participants navigate novel grids on every trial with only knowledge of a few landmarks

1. Learning Phase



Learn about location of landmarks and goal in top-down view of grid

2. Navigation Phase

Start in unlearnt location
Navigate to goal in two ways

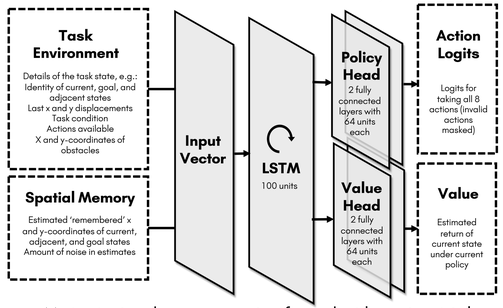
- **'Directions'** – click arrow to move one step in direction (-Vectors)
- **'States'** – click object to move to that state (-Transitions)

Both movement methods allow movement only to **adjacent states** (i.e, they are equivalent but allow dissociation of strategies)

Manipulations: on some blocks, participants may only be allowed to use the 'vectors' only, 'transitions' only, or one side or the other randomly; open field vs cluttered environments

Deep Meta-RL Model

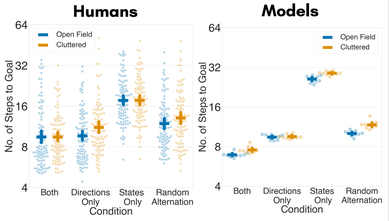
Deep RL model trained with Proximal Policy Optimization



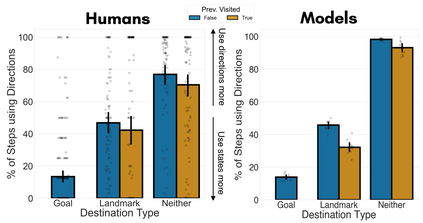
Meta-trained on successive few-shot learning trials
10 random initialization seeds

Human and Model Behavior

Humans and models perform best when allowed to freely choose between response modes



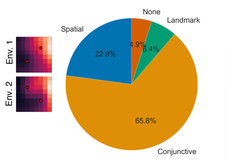
Humans and models rely mostly on directions, but use transitions at **landmarks** and **previously visited states**



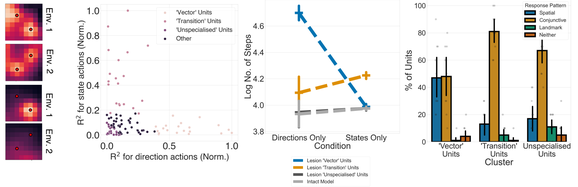
Few-shot planning depends on exploiting generalizable spatial structure with vector-based strategies and sparse transition knowledge with transition-based strategies

Model Representations

Emergence of spatially and landmark-selective units

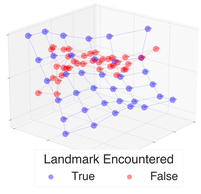


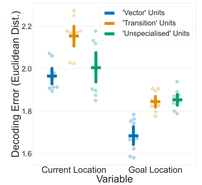
Emergence of 'vector' and 'transition' modules with different characteristic response patterns



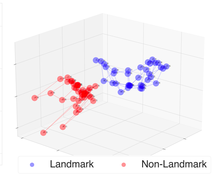
'Vector' and 'transition' units have different representational contents and geometries

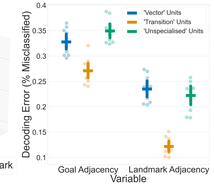
'Vector' Units





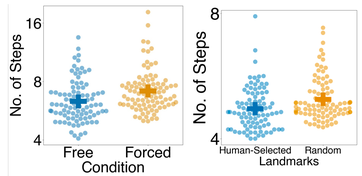
'Transition' Units



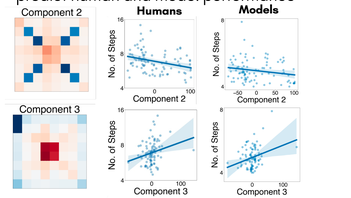


Human Landmark Sampling Behavior

Humans select landmarks that are beneficial for navigation



Principal components of human sampling patterns predict human and model performance



Conclusions

- In both humans and deep RL models, few-shot spatial planning depends on using vectors to head in general goal direction, and transitions near learnt landmarks
- Deep RL models develop 'modules' of units with different response profiles, representational contents, and representational geometries
- Humans are successful in selecting 'landmarks' that are beneficial for few-shot performance

Abstract Link

X: @denislan_