Hierarchy discovery for planning

Momchil

November 7, 2018

1 Motivation

People spontaneously organize their environment into clusters of states (state chunks) which constrain planning [1, 2, 3]. Why and how do we do that?

One normative reason is that hierarchical planning is more efficient in time and memory than "flat" planning [4] (e.g. for BFS, $O(\sqrt[L]{N})$ vs. $O(N)$, where N is the number of states and L is the hierarchy depth). This is consistent with people's limited working memory capacity [5].

But hierarchies are a form of lossy compression, and some hierarchies might result in inefficient planning.

[example]

Solway et al. 2014 [1] provide a formal definition of an "optimal" hierarchy, but they do not specify how the brain might discover it. In this work, we propose a Bayesian cognitive model of hierarchy discovery and show that it is consistent with human behavior.

2 Background

2.1 Preliminaries

We assume a 2-layer hierarchy ($L = 2$). $G = (V, E)$ is the low-level ("flat") graph that is directly observable, where:

- V is the set of vertices (low-level states)
- $E : \{V \times V\}$ is the set of edges

We define $H = (V', E', c, b)$ as the high-level ("H") graph that is not directly observable, where:

- V' is the set of vertices (high-level states, or state chunks)
- $E' : \{V' \times V'\}$ is the set of edges

- $c : V \to V'$ are the state chunk assignments
- $b : E' \to E$ are the bridges

Each low-level state i is assigned to a state chunk $k = c_i$. Each high-level edge (k, l) has a corresponding low-level edge (the bridge) $(i, j) = b_{k,l}$, such that $c_i = k$ and $c_j = l$. For simplicity, we assume both graphs are unweighted and undirected.

2.2 Model

Informally, an algorithm that discovers useful hierarchies would satisfy the following desiderata:

1. Favor smaller state chunks,
2. ...but not too many of them;
3. Favor dense connectivity within chunks,
4. ...and sparse connectivity across chunks,
5. ...with the exception of "bridges" that connect chunks

Intuitively, having too few (e.g. one) or too many state chunks (e.g. each state is its own chunk) creates a degenerate hierarchy that reduces the planning problem to the "flat" scenario, and hence medium-sized chunks are best (desiderata 1 and 2). Additionally, the hierarchy ignores transitions across state chunks, which could lead to suboptimal paths generated by the hierarchical planner. It is therefore best to minimize the number of cross-chunk transitions (desiderata 3 and 4). The exception is bridges, which correspond to the links between state chunks (desideratum 5).

These can be formalized into a generative model for "good" hierarchies:

$$
\begin{aligned}
c &\sim CRP(\alpha) && \text{chunk assignments} \\
p' &\sim Beta(1,1) && \text{H graph density} \\
Pr[(k,l) \in E'] &= p' && \text{H graph edges} \\
Pr[b_{k,l} = (i,j) \mid (k,l) \in E', c_i = k, c_j = l] &= \frac{1}{n_k n_l} && \text{bridges} \\
p &\sim Beta(1,1) && \text{within-chunk density} \\
q &\sim Beta(1,1) && \text{cross-chunk density penalty} \\
Pr[(i,j) \in E \mid c_i = c_j] &= p && \text{within-chunk edges} \\
Pr[(i,j) \in E \mid c_i \neq c_j] &= pq && \text{cross-chunk edges} \\
Pr[(i,j) \in E \mid b_{c_i,c_j} = (i,j)] &= 1 && \text{bridge edges}
\end{aligned}
$$

Where $n_k = |\{i : c_i = k\}|$ is the size of chunk k and CRP is the Chinese restaurant process, a nonparametric prior for clusterings [6].

Hierarchy discovery can then be framed as inverting the generative model (in similar spirit to how k-means clustering can be understood as inference over a Gaussian Mixture model). The posterior probability of H is:

$$\begin{aligned} P(H|G) &\propto P(G|H)P(H) && (1) \\ &= P(E \mid c, b, p, q)P(p)P(q)P(b|E', c)P(E'|p')P(p')P(c) && (2) \end{aligned}$$

2.3 Tasks

Like prevous authors [1, 3], we assume the agent faces a sequence of tasks in G, where each task is to navigate from a starting state s to a goal state g. We assume the agent prefers shorter routes.

Informally, the hierarchy discovery algorithm might account for tasks by chunking together states that frequently co-occur in the same task, since hierarchical planning is optimal within chunks.

Defining $tasks = \{task_t\}$ and $task_t = (s_t, g_t)$, we can augment the generative model as:

$$\begin{aligned} p'' &\sim Beta(1,1) && \text{cross-chunk task penalty} \\ Pr[s_t = i] &= \frac{1}{N} && \text{starting states} \\ Pr[g_t = j \mid s_t = i] &\propto \begin{cases} 1 & \text{if } c_i = c_j \\ p'' & \text{otherwise} \end{cases} && \text{goal states} \end{aligned}$$

Where $N = |V|$ is the total number of states. We denote the observable data as $D = (tasks, G)$. The posterior then becomes:

$$\begin{aligned} P(H|D) &\propto P(D|H)P(H) && (3) \\ &= P(tasks|G, H)P(G|H)P(H) && (4) \\ &= \left[\prod_t P(g_t|s_t, p'', G, H)P(s_t|G, H)\right] P(p'')P(G|H)P(H) && (5) \end{aligned}$$

Where the last two terms are the same as in Eq. 1.

2.4 Inference

We approximate Bayesian inference over H using Metropolis-within-Gibbs sampling [7] (a kind of MCMC) which updates each component of H in turn by sampling from its posterior conditioned on all other components in a single Metropolis-Hastings step. The proposal distribution for continuous components is a Gaussian random walk. The proposal distribution for chunk assignments c_i is the conditional CRP prior (algorithm 5 in [8]).

Our approach can also be interpreted as stochastic hill climbing with respect to a utility function defined by the posterior. This has been previously used to find useful hierarchies for robot navigation [4].

Bibliography & References Cited

[1] Alec Solway, Carlos Diuk, Natalia Córdova, Debbie Yee, Andrew G. Barto, Yael Niv, and Matthew M. Botvinick. Optimal behavioral hierarchy. *PLOS Computational Biology*, 10(8):1–10, 08 2014.

[2] Anna C. Schapiro, Timothy T. Rogers, Natalia I. Cordova, Nicholas B. Turk-Browne, and Matthew M. Botvinick. Neural representations of events arise from temporal community structure. *Nature Neuroscience*, 16(4):486–492, Apr 2013.

[3] Jan Balaguer, Hugo Spiers, Demis Hassabis, and Christopher Summerfield. Neural mechanisms of hierarchical planning in a virtual subway network. *Neuron*, 90(4):893–903, 2016.

[4] Juan A Fernández and Javier González. *Multi-hierarchical representation of large-scale space: Applications to mobile robots*, volume 24. Springer Science & Business Media, 2013.

[5] George A Miller. The magic number seven plus or minus two: Some limits on our capacity for processing information. *Psychological review*, 63:91–97, 1956.

[6] Samuel J Gershman and David M Blei. A tutorial on bayesian nonparametric models. *Journal of Mathematical Psychology*, 56(1):1–12, 2012.

[7] Gareth O Roberts and Jeffrey S Rosenthal. Examples of adaptive mcmc. *Journal of Computational and Graphical Statistics*, 18(2):349–367, 2009.

[8] Radford M Neal. Markov chain sampling methods for dirichlet process mixture models. *Journal of computational and graphical statistics*, 9(2):249–265, 2000.