Predicting Active Cognitive Processes from Task-Based Contrast Images

Dan Birman, Sanmi Koyejo, Russ Poldrack

**Overview**

One of the major goals in neuroimaging is to predict the cognitive processes engaged at a given moment, captured by a brain activation map. This inference is often performed informally based on previous observations of activation in a particular region during a cognitive process (e.g. the amygdala during fear) and reversed by making the assumption that if amygdala activation is observed the participant experienced fear. This reverse inference is problematic; in that we cannot assume that previous research effectively isolated the region’s response as being caused specifically by only a single cognitive process. Closing the bidirectionality of inference is an active research topic and has been previously explored for combinations of cognitive processes, for example entire tasks (Poldrack et al. 2009) and for specific aspects of tasks (Schwartz et al. 2013). It remains unclear whether these predictive models for mapping from brain activation to cognitive processes can be leveraged to clarify our understanding of the processes themselves. Our goal here is to generate predictive models that can help clarify what cognitive processes exist, how they are interrelated, and whether they can be consistently distinguished across tasks. Using inferential models in this way we hope to further our understanding of how neural activity contributes to cognition and our experience of distinct cognitive processes.

**Goals**

The main goal is to predict, from activation maps, the functional cognitive processes—in some sense a description of the task being performed when the activation map was recorded. There are several smaller stepping-stones to achieving this:

1. Forward inference maps for the terms in the task dataset(s). (I.e. when a term is seen, what activation of voxels do we expect)
2. Reverse inference maps for the terms in the task dataset(s). (I.e. what activation map predicts the presence of the term)
3. Leave-one-study out prediction of a study’s terms. (i.e. given an unseen set of study data, but where every process has overlap with another study, return a description of the study’s task)
4. Leave-one-subject out prediction.

**Datasets**

We plan on using task fMRI data from the Open fMRI project (currently 31 studies, many different tasks) and the Human Connectome Project (500 subjects, resting state, tasks: working memory, recognition memory, gambling, motor, language, social cognition, relational processing, emotion). An initial design could train and test on different datasets, or on subsets of both datasets for maximum term coverage. An important constraint is that any term that we include must cover multiple studies.

We plan to use the pre-processed data from both databases as is. Both databases have data in volume formats, on which we can perform our additional analyses.

**Labeling**

The contrasts in the task data will need to be labeled to identify what processes are active at a given moment. We plan to do task labeling by hand, following the cognitive functions outlined by the CognitiveAtlas (<http://cognitiveatlas.org/>). For the stimuli, response, etc, we will follow the outline provided by CogPO (<http://www.wiki.cogpo.org/index.php?title=Main_Page>). The goal is to identify for every experimental condition the structural characteristics of the current task. This might include: stimuli, stimulus modality, instructions, cognitive (implicit?) responses, explicit responses, response modalities, etc. CognitiveAtlas has an excellent database of cognitive functions spanning a wide array of terms. CogPO has poor coverage of cognitive functions but a terse list of stimulus/response modalities.

**Forward Inference**

The process we plan to use is to generate a linear regression model (GLM) for each voxel in the subject activation maps, under the assumption that the response in each voxel, x, is derived from a linear combination of the effects of the task labels on that voxel. As Schwartz et al. point out; by using *term-versus-others* contrasts the GLM formulation estimates each term’s effect while partialling out the effects of the other terms.

**Reverse Inference**

We plan to follow a similar procedure to that outlined in Schwartz et al. 2013 (*2.3 Reverse inference*, pg. 3). In short: we will train one classifier for each term, designed to discriminate the presence or absence of each term label. The combination of multiple classifiers then provides a description of the expected terms in a given brain map.

**References**

Poldrack, R. (2006). Can cognitive processes be inferred from neuroimaging data? *Trends in Cognitive Sciences*, *10*(2), 59–63. doi:10.1016/j.tics.2005.12.004

Poldrack, R. A., Halchenko, Y. O., & Hanson, S. J. (2009). Decoding the Large-Scale Structure of Brain Function by Classifying Mental States Across Individuals. *Psychological Science*, *20*(11), 1364–1372. doi:10.1111/j.1467-9280.2009.02460.x

Schwartz, Y., Thirion, B., & Varoquaux, G. (2013). Mapping paradigm ontologies to and from the brain. In *Advances in Neural Information Processing Systems* (pp. 1673–1681).