Are semantic representations of words radically distributed?

Cox, C. R.¹, Seidenberg, M. S.¹, Binder, J. R. ², Desai, R. H. ² and Rogers, T. R.¹

1. Department of Psychology, University of Wisconsin-Madison
2. Department of Neurology, Medical College of Wisconsin

Introduction

Some theories propose that semantic knowledge about different conceptual domains is modular, that is, encoded in distinct, contiguous, and non-overlapping regions of cortex. [1]

Others propose that semantic knowledge is highly distributed [2]:
- A given region can encode information about more than one conceptual domain.
- Immediate neighbors may encode different types of information.
- Distal regions may be part of the same representation.
- Information encoded by a given region may vary across individuals.

fMRI evidence is often taken to support the modular view; however standard univariate methods for controlling Type 1 error are biased towards discovering modular-looking signal and will fail to discover important signal if the representation is distributed.

We consider a new multivariate approach, LASSO logistic regression, that makes no assumptions about localization of function, and compare the results of this approach to those obtained from a standard univariate contrast analysis of data from a word-reading task.

The dataset

Participants: 16 right handed subjects scanned at the Medical College of Wisconsin.

Stimuli: 1200 letter strings (900 words, 300 nonwords obaying English phonology). 600 words were concrete, 300 abstract. Of the concrete nouns: 94 animals 145 artifacts

Task: Respond to the question “Is it something physical you can experience with your senses?” by button-press for every stimulus.

Procedure: Letter strings presented one at a time at jittered intervals in random order over the course of 10 scanner runs with no repetitions in a rapid event-related design.

Imaging and preprocessing: Participants were scanned in a 3T scanner with a 2-second TR. The first 5 TRs from each run were discarded. Functional data was time-shifted, registered to the anatomy and scaled prior to deconvolution.

Analysis 1: Univariate approach

Deconvolution: Raw BOLD signal was blurred with a 4mm FWHM Gaussian kernel. Multiple linear regression with gamma variate HRF. GLM analysis included single regressors for ANI and ART words. Beta coefficient image for ANI and ART regressors for each subject warped to N27 Talairach space with 3mm³ voxels, and submitted to a within subjects paired t-test.

Corrected for multiple comparisons using FDR (q < 0.05) and establishing a 20-voxel minimum cluster size.

Animals – baseline:

Artifacts – baseline:

Animals – Artifacts:

No Significant Clusters with corrected p in ANI-ART Contrast

Analysis 2: LASSO logistic regression

Univariate analysis failed to find regions selective for animals over artifacts. In analysis 2 we applied L1 (LASSO) regularized logistic regression as a multivariate classifier to identify the same two categories. Classifiers were trained on all cortical voxels for each of 9 individual participants from Analysis 1.

Deconvolution: Multiple linear regression with gamma variate HRF. GLM analysis included a regresor for each word and nonword stimulus, resulting in 1200 volumes with beta coefficients at each voxel for each subject.

Voxel pre-selection: Surface reconstructions for each subject’s anatomy were created and used to select only cortical voxels (i.e. space between pial and white matter ribbons) from the volumes of beta coefficients for each word and subject. All MVPA was performed on cortical selections of beta coefficients.

Classification: For each subject, the logistic LASSO was used as a classifier to identify animals from everything else and artifacts from everything else based on functional data for each word. This classifier minimizes a cost defined by the logistic loss function (i.e., how well the classifier fits the training data) plus the sum of the absolute values of the weights in the classifier:

\[ \beta = \arg \min (1 + e^{-y \cdot \beta}) + \lambda \| \beta \|_1 \]

Cross Validation: The words were separated into 10 blocks, each balanced for number of animals, artifacts, and abstract words. Cross validation was performed at two levels: one to pick a good value of \( \lambda \) from 10 pre-specified choices without peeking, and another to evaluate the chosen \( \lambda \) for all train-test combinations of the 10 blocks of words.

MVPA Results: Above Chance Classification

Aggregate over subjects

MVPA Results: Individual Differences

The role of individual voxels in distributed network varies by subject.

Summary & Conclusions

A massively-univariate analysis fails to identify any regions selectively active for animals or artifacts.

However, the LASSO multivariate classifier is able to pick out animals and artifacts with likelihood greater than chance:
- From a single presentation of a word
- In individual subjects with no smoothing or averaging
- During a task that did not discriminate animals from artifacts.

When considering where the voxels informative to the classifier are located, it is clear why the univariate analysis failed.

- Most voxels encoded information about both animals and artifacts across and often within subjects.
- Those voxels that were category-selective across subjects were sparse and distributed.

We conclude that the semantic representations of animals and artifacts are encoded by patterns of activation over a common set of units and that there are no anatomical regions dedicated to just animal or artifact knowledge. Patterns of activity for animals and artifacts do appear to differ, but they do so by few voxels, in the context of a full distributed network of knowledge, which are highly sparse and spatially distributed.

References

4. Cortical reconstruction and volumetric segmentation was performed with the FreeSurfer image analysis suite, which is documented and freely available for download online (http://surfer.nmr.mgh.harvard.edu/).

Contact information: Chris Cox ccrox@wisc.edu