

aa: Parallel Processing and Efficient Workflows in Matlab on your Cluster or in the Cloud

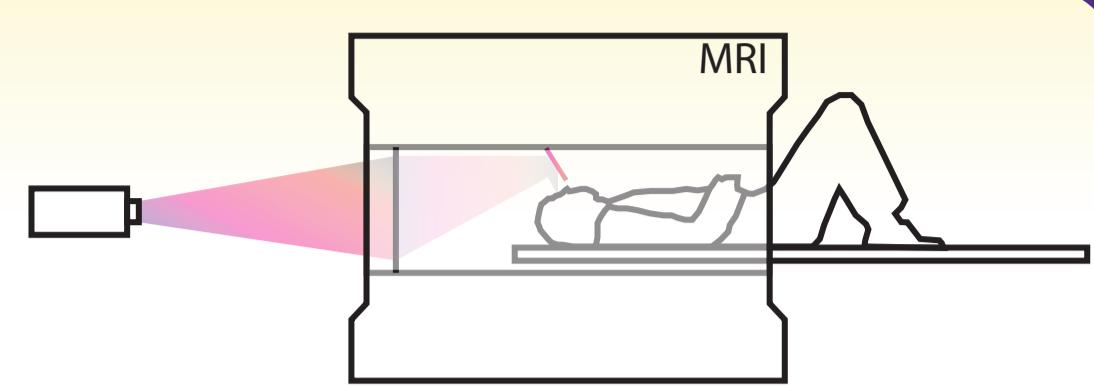
aa4

Rhodri Cusack Brain and Mind Institute, Western University, Canada
Alejandro Vicente-Grabovetsky Donders Institute for Brain, Cognition and Behaviour, The Netherlands
Daniel J. Mitchell MRC Cognition and Brain Sciences Unit, Cambridge United Kingdom
Jonathan E. Peelle Department of Neurology, University of Pennsylvania, United States

automaticanalysis-announce

The challenge

- Number of neuroimaging analysis algorithms
- Algorithm complexity
- Computational demand
- Risk of human error
- Cohort size
- Types of MRI data
- Workflow complexity



What is needed

- | | | |
|---|---|---|
| Efficient computation
<ul style="list-style-type: none"> - Only do what needs to be done - Parallel processing - Cloud enabled | Effective management of analyses
<ul style="list-style-type: none"> - Simple to specify, easy to learn - Keep track of what needs to be done - Good record keeping | Easy and effective code development
<ul style="list-style-type: none"> - D(on't) R(peat) Y(ourself) - Easier to develop within aa than outside - Fully customizable - control everything - Self documenting where possible - github source control |
|---|---|---|

Anatomy of the solution

User script (.m)
Specifies analysis
- Data to process (subjects, sessions)
- Tasklist
- Customises tasks and parameters

Engine (.m)
Controls execution
- Calculates dependencies
- Defines data file structure
- Retrieves data from remote filesystem if necessary
- Calls appropriate modules serially, or in parallel
- Keeps records of what gets run and how long it takes
- Compiles code if necessary
- Supports Condor, GridEngine, Torque

Modules (.m, .xml)
Each module runs a single analysis stage
- Many modules available
- Explicitly defines its inputs and outputs
- Defines domain for execution (i.e., once per study|subject|session)
- Low overhead, immediately reusable
- Data to process passed from engine

Tasklists (.xml)
Describes a series of modules to be executed
- Can include analysis branches
- Modules can be simply reordered

Example user script (.m)

```
% Example aa version 4 user script
% Local processing, no modelling
% Rhodri Cusack, Cambridge Feb 2011

aa_ver4-devel

% Add the directory of this script to the top of
% the matlab path
addpath(fullfile(fileparts(fullfile('fullpath')),'-begin'));

% DEFINE STUDY SPECIFIC PARAMETERS
aap=aarecipe('aap_tasklist_fMRI_typical.xml');
% Directory for analysed data
aap.acq_details.root='/imaging/rhodri/camcan/cc_movie';
% Sub-directory for analysed data
aap.directory_conventions.analysisid='data';

% Define subjects, and EPI series number ordered as
% addsession lines
aap=aas_addsubject(aap,'CBU110000_*',(6));

% One or more sessions
aap=aas_addsession(aap,'movie');

% How many dummies?
aap.acq_details.numdummies=3;

% SITE-SPECIFIC CONFIGURATION
aap=aas_localconfig(aap);

% Do processing
aa_doprocessing(aap);
```

MVPA

Commonly used tasklists & modules

Common

aamod_autoidentifyseries_timtrio

- Scan input DICOM files to get series & acquisitions irrespective of filenames, which are typically site-specific

- Identify structural and fieldmap series numbers

aamod_norm_noss

- Perform normalise & segment with SPM 8

aamod_getdicom_epi

- Find all DICOM files corresponding to the EPI acquisitions

aamod_convert_epi

- Convert the DICOM files to NIFTI format. Handles with multi-echo EPI with various weighting schemes.

aamod_realign

- Perform motion correction with SPM 8

aamod_slicetiming

- Slice timing correction with SPM 8

aamod_coreg_noss

- Co-registers structural to mean EPI using SPM 8

aamod_norm_write

- Apply normalization parameters derived from structural to EPIs

aamod_smooth

- Smooth data

aamod_firstlevel_model

- Runs first level statistical model. Simple specification of events in user script

aamod_firstlevel_contrasts

- Runs first level contrasts. Simple specification of contrasts.

aamod_secondlevel_model

- Runs a t-test across subjects for every first level contrast

aamod_tensor_ica

- Runs individual or group tensor ICA

aamod_meantimecourse

- Calculates mean time course for each voxel across subjects

aamod_moviecorr_meantimecourse

- Calculates correlation of each subject's timecourse with mean

aamod_moviecorr_summary

- Statistics to find which correlations are significant across subjects

aamod_biascorrect_segment8

- Run New Segment (SPM 8) and save bias-corrected image for segmenting

aamod_segment8

- Tissue class segmentation using New Segment (SPM 8)

aamod_structuralstats

- Retrieve total tissue class volume and TIV from segmented images

aamod_darTEL_createmplate

- Use DARTEL to create a template

aamod_darTEL_normmmi

- Write DARTEL-warped images to MNI space

aamod_normalisebytotalgrey

- Divide segmented images by total grey matter (proportional scaling)

aamod_MVPaa_brain_1st

- Runs an MVPA searchlight on a set of beta or T-values (typically in native space)

aamod_MVPaa_brain_SPM

- Convert results from searchlight into nifti images readable in SPM

aamod_unnormalise_roi

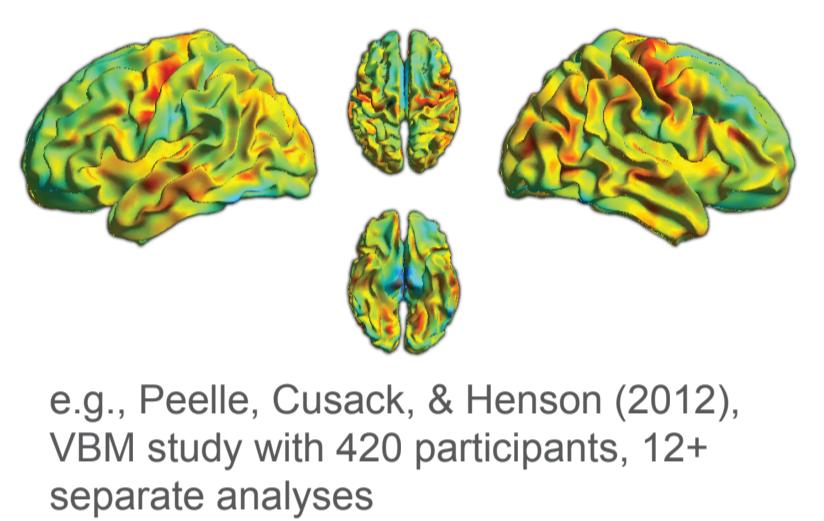
- Set ROIs from standard space into subject space

aamod_MVPaa_roi_1st

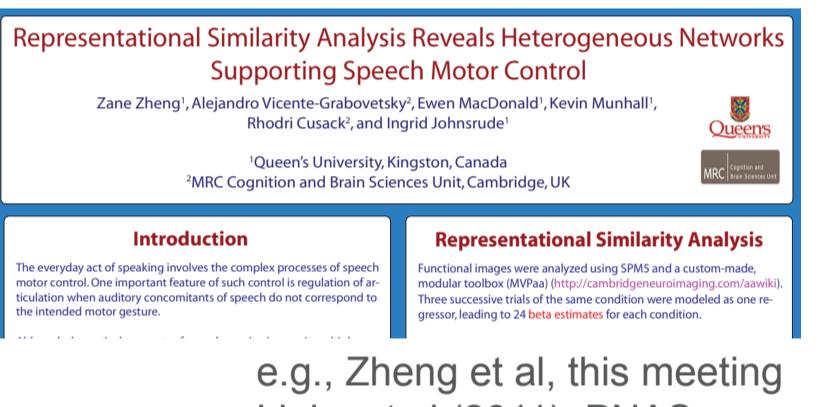
- Runs an MVPA analysis within an ROI, using a set of beta or T-values (typically in native space)



e.g., Ongoing studies by www.camcan.org



e.g., Peelle, Cusack, & Henson (2012), VBM study with 420 participants, 12+ separate analyses



e.g., Zheng et al., this meeting Linke et al (2011), PNAS

Example tasklist (.xml)

```
<?xml version="1.0" encoding="utf-8"?>
<aap>
  <tasklist>
    <initialisation>
      <module><name>aamod_checkparameters</name></module>
      <module><name>aamod_evaluatesubjectnames</name></module>
      <module><name>aamod_study_init</name></module>
      <module><name>aamod_news subj_init</name></module>
    </initialisation>

    <main>
      <module><name>aamod_autoidentifyseries_timtrio</name></module>
      <module><name>aamod_get_dicom_structural</name></module>
      <module><name>aamod_get_dicom_epi</name></module>
      <module><name>aamod_convertmaps</name></module>
      <module><name>aamod_copystructural</name></module>
      <module><name>aamod_convert_epis</name></module>
      <module><name>aamod_realign</name></module>
      <module><name>aamod_tsdfaffana</name></module>
      <module><name>aamod_slicetiming</name></module>
      <module><name>aamod_coreg_noss</name></module>
      <module><name>aamod_norm_noss</name></module>
      <module><name>aamod_norm_write</name></module>
      <module><name>aamod_smooth</name></module>
    </main>
  </tasklist>
</aap>
```

The elegance of streams

Modules explicitly define their inputs and outputs
e.g., fragment from aamod_realign.xml

```
<inputstreams>
  <stream>epi</stream>
</inputstreams>
<outputstreams>
  <stream>epi</stream>
  <stream>meanepi</stream>
</outputstreams>
```

Takes a set of EPI volumes; produces another set of EPI volumes, and a mean EPI volume

This allows calculation of the dependency map - the flow of data amongst the modules, permitting:
- Engine to schedule concurrent (parallel) processing of modules that are not dependent on each other
- Branched analyses in which a tasklist splits and is completed in multiple ways
- Engine can retrieval/archive data from/to a remote server & convert formats
- No need for file filtering by prefix simplifying code in modules and allowing simple reordering of modules in a tasklist
- Engine can calculate which modules need to be redone when some given module is changed and rerun

Track record

Mature platform

- 8 years old, on fourth major release and under active development by an international team
- 50+ users have analyzed 100+ studies comprising 1000s of subjects and produced scores of publications
- Has been used to analyze MRI (fMRI/VBM/MVPA) & MEG data, for example:

Cusack, Rhodri, Mitchell, D. J., & Duncan, J. (2009). 'Discrete Object Representation, Attention Switching, and Task Difficulty in the Parietal Lobe.' Journal of Cognitive Neuroscience. doi:10.1162/jocn.2009.21194
Linke, A. C., Vicente-Grabovetsky, A., & Cusack, R. (2011). 'Stimulus-specific suppression preserves information in auditory short-term memory.' Proceedings of the National Academy of Sciences of the United States of America, 108(31). doi:10.1073/pnas.1102118108
Peelle, J. E., Cusack, R., & Henson, R. N. A. (2012). 'Adjusting for global effects in voxel-based morphometry: Gray matter decline in normal aging.' NeuroImage 60:1503-1516.

- Tested on larger cohorts (N=100-700)
- Tested on Condor, Torque, GridEngine
- Tested at many sites
- Cloud enabled

How to get it

aa may be downloaded from github.com/rhodricusack/automaticanalysis
Google group discussion list is here: automaticanalysis-announce