NEURAL MAPPING OF MATHEMATICAL ACTIVITY

FRANK QUINN

ABSTRACT. This is a proposal for detailed mapping of neural implementation of elementary mathematical functionality. Mathematics is a good candidate because it is a learned activity that is highly structured and can provide detailed external work records. Carrying out the proposal will still require refined neuroscience techniques, particularly in magnetoencephalography.

CONTENTS

1. INTRODUCTION

We might ask about any cognitive activity:

- What are the locations and functions of component neural processes?
- Which processes communicate with which others, and what is the timing and function of these communications?
- How is the activity organized and controlled?

These are sweeping questions, but the mathematical context has powerful advantages including: unambiguous criteria for correct function; non-neural data divides the activity into small subtasks; and mathematical structure puts strong constraints on how these subtasks might be subdivided and organized for neural implementation. The consequence, roughly, is that if neuroscience can answer the 'when' and 'where' questions then functionality can be inferred from the constraints and other data.

Date: Draft, April 2011.

Roughly, the plan is to record external memory use (scratch work) and attention focus (eye movement) while subjects are doing elementary mathematical work. The neuroscience tasks, guided by this data, are:

- determine which neural regions are active at any given time during the task,
- identify source, target, and timing of communications between these regions.

More precise versions are given in the body of the proposal.

These questions certainly push the limits of current technology, but perhaps not beyond reach. The proposal describes a bootstrap approach using fMRI, DSI, EEG, MEG, etc. The bottleneck is likely to be determination of the destination of an inter-region communication whose source and timing are known, using MEG.

2. External data

The proposal is to exploit special features of mathematical work to map its neural implementation. Use of external working memory and eye movement to organize the work is illustrated with an example in §2.1, 2.2. Consequences are that subtasks can be studied *in situ* rather than in artificial isolation; statistical information can be obtained by comparing equivalent subtasks of different tasks; and overall complexity of the task does not have to be controlled.

2.1. **Sample task.**

This illustrates the idea with the task-separated version of polynomial multiplication discussed in [9] §2.1.2. The first point is that external memory gives a detailed record of the nature and timing of the more computational subtasks.

Example: find the coefficient on x^3 when

$$(5x^3 + 3x^2 - x + 5a)(x^3 + (2-a)x^2 - a)$$

is written as a polynomial in x.

The first step is to collect pairs of coefficients (one from each factor) whose total x coefficient is 3. The first term in the first factor has exponent 3 and coefficient 5. The 5 is recorded in scratch work, and the second factor is scanned for the complementary exponent (0). The coefficient on x^0 is $-a$, so this is recorded in scratch work. Now repeat for the next term in the first factor. The coefficient is 3, and this is recorded in scratch work, and the exponent is 2. Now scan the second factor for the complementary exponent (1). There are no such terms so coefficient 0 is entered in scratch work. At this point the scratch work should look like:

$$(5)(-a) + (3)(0) +$$

From this we can infer that the next act will be to scan the first factor for terms with exponent 1, and record the coefficient. Moreover the coefficients are being read purely as strings of symbols to be copied, because no manipulations are being done.

After the collection step individual multiplications are done. The first is $(5)(-a)$ which is recorded in scratch work in standard form $-5a$:

$$\underbrace{(5)(-a)}_{-5a} + (3)(0) + (-1)(2-a) + (5a)(1)$$

Note that the components are in standard locations so scanning is less intrusive than in the first step. At this point we know that the next act will be to read

3 and 0 as numbers, multiply them, and record the outcome. The result of the multiplication step is:

$$\underbrace{(5)(-a)}_{-5a}+\underbrace{(3)(0)}_{0}+\underbrace{(-1)(2-a)}_{a-2}+\underbrace{(5a)(1)}_{5a}$$

Finally the additions are done. Partway through, the scratch work might show:

$$\underbrace{\underbrace{(5)(-a)}_{-5a}+\underbrace{(3)(0)}_{0}+\underbrace{(-1)(2-a)}_{a-2}+\underbrace{(5a)(1)}_{5a}}_{a}$$

This would indicate that the a terms were identified and their coefficients added. The next act will be to add the constant terms.

2.2. **Eye movement and organizational structure.** External working memory gives good access to computational activity because steps are recorded in some detail, and input is usually in standard position in the visual field. Organizational activities are more varied and leave fewer clues. Some examples:

- Initial preparation probably includes forming tree-like representations of structure of compound expressions. These representations may associate significance to specific locations rather than internalizing details.
- Activities such as collecting coefficients in a product require locating relevant pieces (the scanning described above), and copying to standard formats for computational steps.

Tracking eye movement will provide a record of locations of attention focus that can be used to divide organizational activities into subtasks.

3. General experimental design

Most detailed neural studies use tasks that seem simple, (finger tapping, responding to a stimulus) but they call on systems that have been integrated and optimized by evolutionary pressure. Activity is fast, tightly coordinated, and largely short-range ("small world" structure). In most cases this will be very difficult to understand in any detail.

Mathematics is an abstract learned activity, and likely to be strongly modular in the sense that most subtasks should call on a relatively small number of regions. Moreover, the functionality was developed for other purposes and certainly not optimized for this use. As a result the regions should usually be physically separated and inter-region communications should use generic and physically long neural connections. Finally, inter-region communication should be relatively sparse: physical length probably limits the repetition rate, and makes the signals vulnerable to crosstalk and timing problems. We clarify the goals, assuming this qualitative picture is valid. Neuroscience goals are:

- "Primitive" processes are those located in a specific neural region. The goal is to determine the regions used by each subtask and, for timing purposes, track activity in these regions. We do not expect to understand activity in a single region.
- Assuming sparseness, we would like to detect essentially all long-range communications originating in a region of interest.

- If such a communication is addressed to another region of interest, identify this region.
- Identify Input/Output (I/O) interaction with external memory. Details will be complex and unimportant, so for these purposes the full visual and motor areas can probably be considered single "regions".

The plan, again, is to correlate this data with eye movement and the external-memory record, and from this infer the jobs done by the individual regions, the essential information content of inter-regional communication, etc.

3.1. **Experimental subjects.** These studies must be done with single individuals. Studies with multiple individuals depend on commonality that can hardly be expected at the target level of detail. This may change after the methodology is well-developed, and after individual variation has been quantified through single-subject studies.

Many of the techniques require statistical amalgamation of multiple runs with essentially identical activity. This further constrains subject selection:

- the subject should be experienced and fully literate, so patterns are well-established and will not change over time (i.e. no learning).
- the subject should be sufficiently disciplined to repeat elementary operations dozens of times without losing focus, skipping steps, etc.
- performance errors seriously disrupt patterns [3, 11, 4]. Data segments containing errors must be removed from the record (for the baseline study, at least), so high accuracy is needed.

4. Detailed design

This section suggests how to go about getting the data needed. The approach is a bootstrap, with the goal at each stage is to characterize features of interest sufficiently well that they are accessible to study at the next stage.

4.1. **fMRI and regions of interest.** The first neuroscience step is fMRI imaging of the subject working a sequence of sample problems.

- This could be calibrated by first imaging the "default network". Here this would mean thinking of nothing in particular, and writing random symbols and numbers to provide baselines for input/output activity in external memory.
- Previous work, e.g. [8, 10], suggests 5–10 regions of interest will be involved in sub-activities identified by external data. The actual number could easily be twice this because variation is much better controlled in this study. The total number is not a problem as long as there are a significant number of subtasks that involve no more than six or seven regions.
- Previous work suggests that the input/output associated with external memory may be sufficiently automatic and effortless that it will not provoke significant or localized BOLD signal in visual or motor areas. These still need to be included as possible source or targets of interprocess communications.

"Network modeling" is often done using only fMRI data. This is harder than it looks [12], and even the best results fall far short of what is needed here.

4.2. **EEG, MEG and regional activity.** The input for this stage is EEG and MEG data for repetitions of trials used in the fMRI step. It is possible to collect simultaneous fMRI and EEG data [7], and it may someday be possible to collect simultaneous MEG data as well [2]. However the technical challenges are enormous and satisfactory sensitivity is a long way off. For the foreseeable future we have to rely on consistency across trials, and collate data using external activity.

The basic information is in the EEG record, and the function of MEG at this point is to sharpen the picture. The idea is that "computations" are done by neurons in the grey matter on the cortical surface, and communications between groups of neurons goes along fibers in white matter that are (at their ends) roughly perpendicular to the surface. Such communications generate Event-Related Potentials (ERP) in EEG data. Fitting MEG data to a lattice of dipoles located perpendicular to the cortical surface identifies magnetic Event-Related Fields (ERF). Collating these data improves on both [1, 6]

The goals are high-precision timing of activity in each region, and detection of initiation of inter-region communications. The latter should be sparse, complex, energetic, and usually near the end of regional activity, so it should be possible to distinguish them from intra-regional signals.

The ideal goal is real-time detection of these events in *individual trials*. The step described below requires statistical combination of more delicate measurements, and is more likely to succeed if event timing is known exactly. Vigorous correction for artifacts will be necessary, and there may be periods when EKG, eye-movement, or other artifacts are too strong for correction. Special sensor placement or design may help. It should be possible to roughly anticipate timing and firing sequences through statistical comparison of equivalent subtasks, and this narrowing of focus may improve detection. If single-event detection proves impossible then the goal is very sharp statistical data.

4.3. **MEG and inter-region communication.** The final task is to determine the target of an inter-region signal, when source and timing are determined as in the previous section, and if that the target is among the currently active regions of interest or an I/O area. The hope is that comparison with a library of ERF patterns can extract this from magnetic field data.

These patterns should be dynamic (time-dependent):

- Transmission time is long (4-20ms) compared to propagation of fields and sensor responses.
- Indirect paths should have complex dynamic signatures, and these might be easier to identify and characterize than static measurements.
- Source effects are likely to dominate early dynamic structure, and endpoint influences should be stronger in the tail.

It should be possible to get some of these ERF patterns directly. The onset of ERP spikes can be used as gates to trigger high-resolution MEG data collection. After that, correlation of external memory, EEG patterns, etc. should identify events in different parts of the record but likely to be functionally equivalent. In some cases it should be possible to identify the target region, or narrow it to a few possibilities. Statistically combine MEG data for these events to obtain a partial library.

For the most part, comparison ERF patterns will have to come from computed models. This requires determination of anatomical connections between the regions of interest, which can be done using Diffusion Spectrum Imaging (DSI, [5]). Generic connectivity maps may be satisfactory once patterns and methodology are established, but errors in identifying regions or connections will significantly hinder development [12].

Even when the paths are reasonably well characterized, this modeling faces enormous difficulties due to path irregularity, uncertainties about *in vivo* conduction characteristics, and field distortion and attenuation by anatomical structures. The empirical partial library of ERF patterns mentioned above will be essential for developing models.

5. Conclusion

If this mapping scheme can be carried out it would be sure to reveal surprises. Current "connectivity" studies describe temporal correlations rather than interprocess communication. It is much more likely that a few regions organize the others, and there is relatively little direct communication between subsidiary processes. We may also find that the fMRI BOLD data does not quite track activity, and obvious gaps may lead to regions that play important roles but don't work hard enough to make it through current fMRI data filters. This could improve the filters. Finally, there will be surprises in the explicit regional functionalities revealed this way.

References

[1] Babiloni, F; Mattia, D; Babiloni, C; Astolfi, L; Salinari, S; Basilisco, A; Rossini, P.M; Marciani, M.G; Cincotti, F; *Multimodal integration of EEG, MEG and fMRI data for the solution of the neuroimage puzzle*, Magn. Reson. Imaging 22 (2004) 1471–1476.

[2] Burghoff, Martin, et. al. *On the feasibility of neurocurrent imaging by low-field nuclear magnetic resonance*, Applied Physics Letters **96**, 233701 (2010).

[3] Eichele H; Juvodden H. T; Ullsperger M; Eichele T; *Mal-adaptation of event-related EEG responses preceding performance errors*, Frontiers of Human Neurosci. **4** (2010) 65.

[4] Fleming, S.; Weil, R.; Nagy, Z.; Dolan, R.; Rees, G.; *Relating introspective accuracy to individual differences in brain structure*, Science **329** (2010) pp. 1541–1543.

[5] Hagmann P; Cammoun L; Gigandet X; Meuli R; Honey CJ; Wedeen, VJ; Sporns, O; *Mapping the structural core of human cerebral cortex* PLoS Biol **6**(7) (2008) pp. 1479-1493.

[6] Henson R.N; Mouchlianitis, Elias; Friston, Karl J; *MEG and EEG data fusion: Simultaneous localisation of face-evoked responses* Neuroimage. 47(2) (2009) 581–589.

[7] Herrmann, Christoph S.; Debener, Stefan, *Integration of EEG and fMRI Simultaneous recording of EEG and BOLD responses: A historical perspective*, International Journal of Psychophysiology **67** (2008) pp. 161–168.

[8] Krueger F., Landgraf S., van der Meer E., Deshpande G., Hu X, *Effective connectivity of the multiplication network: A functional MRI and multivariate Granger causality mapping study*, Human Brain Mapping (Wiley), 2010.

[9] Quinn, F; *Neuroscience experiments for mathematics education* preprint Jan 2011.

[10] Krueger, F.A.; Spampinato, M. V.; Pardini, M.; Pajevic, S.; Wood, J. N.; Weiss, G. H.; Landgraf, S.; Grafman, J.; *Integral calculus problem solving: An fMRI investigation*, Neuroreport **19** (2008) 1095–1099.

[11] Ravizza, S. M.; Anderson, J. R.; Carter, C. S. *Errors of mathematical processing: The relationship of accuracy to neural regions associated with retrieval or representation of the problem state* Brain Research, **1238** (2008) pp. 118–126.

[12] Smith, S.M; Miller, K. L; Salimi-Khorshidi, G.; Webster, M; Beckmann, C. F; Nichols, T. E; Ramsey, J.D; Woolrich, M. W.*Network modelling methods for FMRI* NeuroImage **54** (2011) 875891