Modeling & Analysis
WE 261

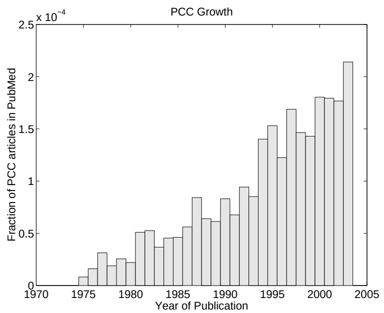
Mining Posterior Cingulate

Finn Årup Nielsen[1,2], Daniela Balslev[1,3], Lars Kai Hansen[2]

[1]Neurobiology Research Unit, Rigshospitalet, Copenhagen, Denmark.
[2]Informatics and Mathematical Modelling, Technical University of Denmark, Lyngby, Denmark
[3]Danish Research Centre for Magnetic Resonance, Hvidovre Hospital, Hvidovre, Denmark

Abstract

We devise a general method for automatic meta-analyses in neuroscience and apply it on text data from published functional imaging studies to extract main functions associated with a brain area — the posterior cingulate cortex. Abstracts from PubMed are downloaded, words extracted and converted to a bag-of-words matrix representation. The combined data is analyzed with hierarchical non-negative matrix factorization. We find that the prominent themes in the PCC corpus are episodic memory retrieval and pain. We further characterize the distribution in PCC of the Talairach coordinates available in some of the articles. This shows a tendency to functional segregation between memory and pain components where memory activations are predominantly in the caudal part and pain in the rostral part of PCC.



Number of posterior cingulate entries in the PubMed database as a function of the year of publication. The query was “("posterior cingulate" OR "posterior cingulum" OR "retrosplenial" OR "retrosplenium")”.

Download of PubMed Abstracts

We download abstracts from the PubMed Web-service by restricting the search to posterior cingulate area and functional neuroimaging with the following query:

```
("posterior cingulate" OR
 "posterior cingulum" OR
  "retrosplenial" OR
  "retrosplenium") AND
("magnetic resonance imaging" OR
 "positron emission tomography")
```

Vectorization of Abstracts

Vectorization of the abstracts using “bag-of-words” and an extensive stop-word list. The result is an abstract × term matrix $\mathbf{X}(N \times P)$.

Text Clustering with NMF

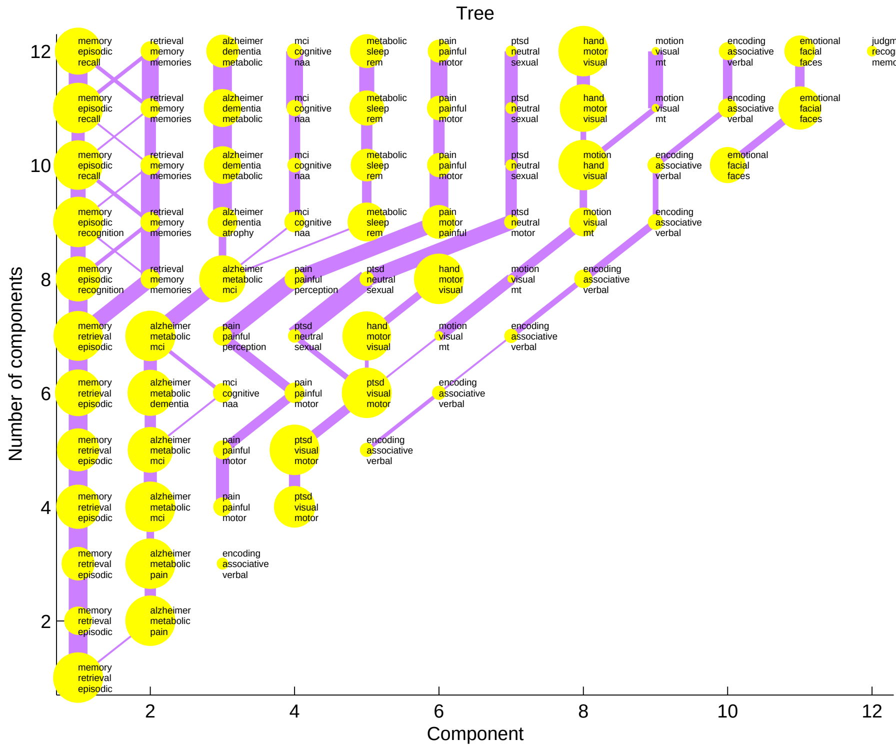
Non-negative matrix factorization (NMF) [1] is use the cluster the vectorized abstracts

$$\mathbf{X} = \mathbf{W}\mathbf{H} + \mathbf{E}, \quad (1)$$

with $\mathbf{W}(N \times K)$, $\mathbf{H}(K \times P)$, K varied between $1 \ldots 12$ and the cost function E is defined as

$$E = \text{trace}(\mathbf{E}\mathbf{E}^\mathsf{T}) = ||\mathbf{E}||_F^2, \quad (2)$$

Hierachical NMF Clustering on Posterior Cingulate PubMed Abstracts



NMF tree with components from non-negative matrix factorization of the bag-of-words converted posterior cingulate abstracts. The nodes indicated by yellow dots represent each a specific component $k \in \{1 \ldots K\}$ for a specific K. The y-axis is indicating a specific number of classes K. The size of the dots indicates the fraction of documents assigned to the component.

The major components appear as memory, Alzheimer’s disease, and pain.

The prominent “memory” component is in alignment with the major review of 275 functional neuroimaging studies [2] that found “episodic memory retrieval” to be associated with posterior cingulate.

The Alzheimer’s disease topic appears due to studies showing hypoperfusion in the posterior cingulate cortex, e.g., [3].

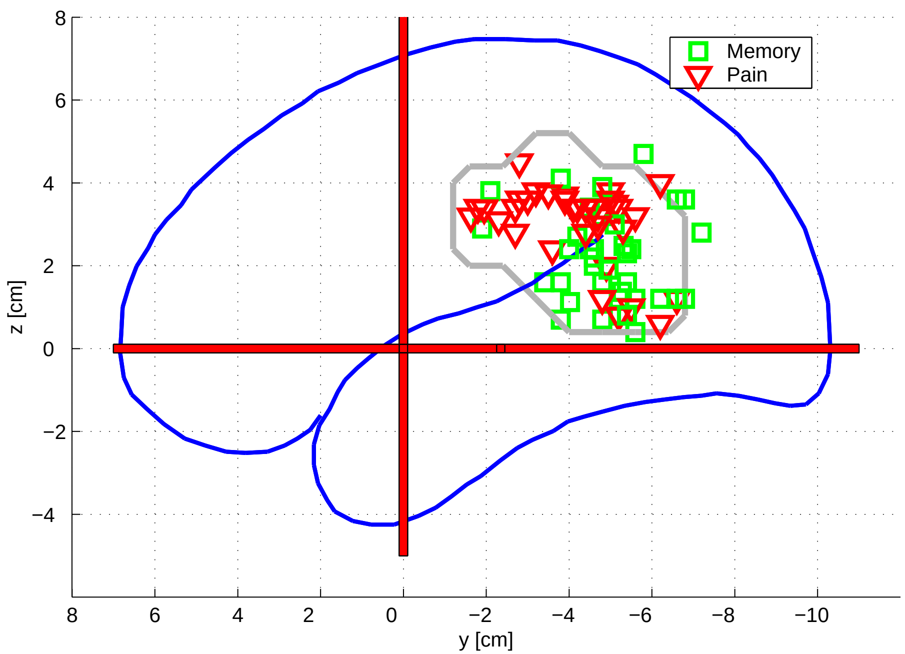
Memory component

#	Load	Title	Ref
1	7.81	Remembering the past: two facets of episodic memory explored with positron emission tomography.	[4]
2	7.56	Differential remoteness and emotional tone modulate the neural correlates of autobiographical memory.	[5]
3	5.74	Differential modulation of a common memory retrieval network revealed by positron emission tomography.	[6]
4	5.11	Remembering familiar people: the posterior cingulate cortex and autobiographical memory retrieval.	[7]
5	5.06	Neuroanatomical correlates of episodic encoding and retrieval in young and elderly subjects.	[8]
6	4.89	Verbal encoding deficits in a patient with a left retrosplenial lesion.	[9]
7	4.73	The functional neuroanatomy of episodic memory: the role of the frontal lobes, the hippocampal formation, and other areas.	[10]
		⋮	

Memory class: Abstracts associated with the memory class sorted according to loading on the NMF component, i.e., values in $\mathbf{w}_{k=\text{memory}}$. For this specific list the $(1,4)$-node from NMF tree figure was selected. The first 7 items in the list all investigate some aspect of memory.

Similar lists can be constructed for Alzheimer's disease and pain. The major part of these studies in the top of the lists is directly related to Alzheimer's disease and pain.

Distribution of Talairach Coordinates from Memory and Pain articles



Distribution of memory and pain brain activations in the posterior cingulate cortex shown on a sagittal plot: y is the AP-axis with posterior as negative. The blue outline follows that of the Talairach atlas. The gray outline is an isocurvature in a probability volume for posterior cingulate cortex based on modeling of coordinates from the Brede database. Green squares are associated with "memory" articles and red triangles with "pain" articles.

There is functional heterogeneity in the posterior cingulate since the foci of the memory and pain activations are different: Memory in the caudal part and pain in the rostral part ($P \approx 0.002$, Hotelling's T^2 test and a Mahalanobis permutation test).

Automated Neuroinformatic Data Mining

We have presented a method for discovering major themes for a specific brain area. Apart from the manual entering of Talairach coordinates our method can be completely automatized. Our method is not confined to the analysis of the posterior cingulate but can be applied to other brain regions or other words and phrases and their relation to brain anatomy can be identified.

Availability and Acknowledgment

The stop word lists and the tools for the analysis are available in the Brede neuroinformatics toolbox [11] presently available from http://hendrix.imm.dtu.dk/software/brede/.

Finn Årup Nielsen is funded by the Villum Kann Rasmussen Foundation.

References

[1] Lee DD and Seung HS. Learning the parts of objects by non-negative matrix factorization. *Nature*, 1999; 401:788–791.

[2] Cabeza R and Nyberg L. Imaging cognition II: An empirical review of 275 PET and fMRI studies. *Journal of Cognitive Neuroscience*, 2000;12:1–47.

[3] Minoshima S, Foster NL, and Kuhl DE. Posterior cingulate cortex in Alzheimer's disease. *The Lancet*, 1994; 344:895.

[4] Andreasen NC, et al. Remembering the past: two facets of episodic memory explored with positron emission tomography. *American Journal of Psychiatry*, 1995; 152:1576–1585.

[5] Piefke M, et al. Differential remoteness and emotional tone modulate the neural correlates of autobiographical memory. *Brain*, 2003;126:650–668.

[6] Maguire EA and Mummery CJ. Differential modulation of a common memory retrieval network revealed by positron emission tomography. *Hippocampus*, 1999; 9:54–61.

[7] Maddock RJ, Garrett AS, and Buonocore MH. Remembering familiar people: the posterior cingulate cortex and autobiographical memory retrieval. *Neuroscience*, 2001;104:667–676.

[8] Daselaar SM, et al. Neuroanatomical correlates of episodic encoding and retrieval in young and elderly subjects. *Brain*, 2003;126:43–56.

[9] McDonald CR, et al. Verbal encoding deficits in a patient with a left retrosplenial lesion. *Neurocase*, 2001; 7:405–417.

[10] Desgranges B, Baron JC, and Eustache F. The functional neuroanatomy of episodic memory: the role of the frontal lobes, the hippocampal formation, and other areas. *NeuroImage*, 1998;8:198–213.

[11] Nielsen FÅ and Hansen LK. Experiences with Matlab and VRML in functional neuroimaging visualizations. In Klasky S and Thorpe S, eds., *VDE2000 - Visualization Development Environments, Workshop Proceedings, Princeton, New Jersey, USA, April 27–28, 2000*. Princeton Plasma Physics Laboratory, Princeton, New Jersey, 2000; pages 76–81.