AN INFORMATION THEORETIC REPRESENTATION OF HUMAN BRAIN FOR DECODING MENTAL STATES OF COMPLEX PROBLEM SOLVING

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF INFORMATICS OF
THE MIDDLE EAST TECHNICAL UNIVERSITY
BY

GONUL GUNAL DEGIRMENDERELI

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE
OF
DOCTOR OF PHILOSOPHY
IN
THE DEPARTMENT OF COGNITIVE SCIENCES

FEBRUARY 2022

AN INFORMATION THEORETIC REPRESENTATION OF HUMAN BRAIN FOR DECODING MENTAL STATES OF COMPLEX PROBLEM SOLVING

submitted by **GONUL GUNAL DEGIRMENDERELI** in partial fulfillment of the requirements for the degree of **Doctor of Philosophy in Cognitive Sciences Department, Middle East Technical University** by,

Date: 07.02.2022

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Surname: GONUL GUNAL DEGIRMENDERELI

Signature :

ABSTRACT

AN INFORMATION THEORETIC REPRESENTATION OF HUMAN BRAIN FOR DECODING MENTAL STATES OF COMPLEX PROBLEM SOLVING

DEGIRMENDERELI, GONUL GUNAL
Ph.D., Department of Cognitive Sciences
Supervisor: Prof. Dr. Fatos T. YARMAN VURAL

February 2022, 133 pages

In this thesis, we propose an information theoretic method for the representation of human brain activity to decode mental states of a high-order cognitive process, complex problem solving (CPS) using functional magnetic resonance images.

First, we aim to identify the active regions and represent underlying cognitive states by measuring the information content of anatomical regions for expert and novice problem solvers during the main phases of problem solving, namely planning and execution. Based on Shannon's entropy definition, we define two new measures: static and dynamic entropy. We investigate the relationship between problem solving phases and the entropy values of anatomical regions. The defined entropy measures successfully identify active brain regions involved in complex problem solving. Anatomical regions with low entropy are consistent with active regions recognized by experimental neuroscience.

Then, we introduce a novel method to estimate static and dynamic brain networks using Kulback-Leibler divergence (relative entropy) for representing the complex problem solving task. We investigate the validity of the estimated brain networks by modeling the planning and execution phases of the complex problem solving. The suggested computational network model is tested by a classification algorithm to discriminate the two phases of complex problem solving. It is observed that, the suggested computational models successfully discriminate the planning and execution phases of the complex problem solving with more than 90% accuracy.

Our results show that the proposed entropy and relative entropy measures hold strong promise for identifying active regions, detecting mind states and predicting brain networks associated with complex problem solving.

Keywords: Brain Decoding, Shannon Entropy, Kullback-Leibler Divergence, Brain Networks, Complex Problem Solving (CPS)

ÖZ

KARMAŞIK PROBLEM ÇÖZMENİN ZİHİNSEL DURUMLARININ BİLGİ TEORİSİYLE TEMSİLİ

DEGIRMENDERELI, GONUL GUNAL
Doktora, Bilişsel Bilimler Bölümü
Tez Yöneticisi: Prof. Dr. Fatos T. YARMAN VURAL

Şubat 2022, 133 sayfa

Bu tezde, üst seviye bir bilişsel süreç olan karmaşık problem çözme ile ilgili zihinsel durumları çözümlemek amacıyla, insan beyin aktivitesinin temsili için Shannon bilgi teorisine dayalı bir yöntem önerilmektedir. Bu yöntem ile, karmaşık problem çözme görevindeki bilişsel durumların hesaplamalı modelleri oluşturulmuştur.

İlk aşamada, problem çözüm süreci ile beyin bölgelerinin entropi değerleri arasındaki ilişki araştırılmıştır. Karmaşık problem çözme sürecinde beyin bölgelerinin içerdikleri bilgi miktarları ölçülerek, aktif bölgelerin belirlenmesi ve başarılı ve başarısız problem çözücüler için bu sürecin temel aşamalarındaki zihin durumlarının tanımlanması hedeflenmiştir. Bu amaçla Shannon Enformasyon Kuramı temel alınarak, insan beyni için iki yeni tanım geliştirilmiştir. Statik ve dinamik entropi adını verdiğimiz bu tanımlar kullanılarak anatomik bölgeler ve bunlar arasındaki ilişkiler modellenmiş, önerilen entropi ölçümlerinin problem çözme sürecindeki aktif beyin bölgelerini başarıyla tanımladığı görülmüştür. Problem çözme sürecinde düşük entropiye sahip anatomik bölgeler literatürde deneysel sinirbilim tarafından tanımlanan aktif bölgelerle uyumludur.

Ardından, göreceli entropi olarak adlandırılan Kullback-Leibler sapması yöntemi ile, karmaşık problem çözme sırasında beyin bölgeleri arasındaki ilişki ölçülerek, başarılı ve başarısız problem çözücüler için bu sürecin temel aşamalarındaki fonksiyonel beyin ağları tahminlenmiştir. Önerilen beyin ağı modeli fonksiyonel Manyetik Rezonans Görüntüleme (fMRG) verileri üzerinde Destek Vektör Makineleri (SVM) algo-

ritması kullanarak test edilmiş, planlama ve yürütme fazlarını %90'ın üzerinde başarı ile sınıflandırmıştır.

Önerilen hesaplama modellerinin, karmaşık problem çözme sürecindeki aktif beyin bölgelerini saptamak, statik ve dinamik beyin ağlarını tahmin etmek, ve zihin durumlarını ortaya çıkarmak için umut verici olduğu görülmüştür.

Anahtar Kelimeler: Beyin Çözümleme, Shannon Entropi, Kullback-Leibler Sapması, Beyin Ağları, Karmaşık Problem Çözme (KPÇ)

To all my teachers who made me experience the joy of learning and understanding

ACKNOWLEDGMENTS

I would like to express my heartfelt thanks and deepest gratitude to my advisor Prof. Dr. Fatos Yarman Vural for her invaluable guidance with her extensive knowledge, experience, continuous support, encouragement, and patience during my Ph.D. study. She believed in my potential, patiently evaluated my work, gave constructive feedback and guided me to bring out my best. Most importantly, she helped me to comprehend the scientific thinking and approach, allowed me to advance my scientific understanding, and pushed my boundaries to progress on the path of being a good researcher. It was a great opportunity and honor for me to have her guide me throughout my thesis studies. She not only became an advisor but, also, a very valuable friend and source of inspiration for me in all aspects of life with her personality, vision, wisdom of life, enthusiasm to teach and help, efforts to bring the best out of her environment, and her approach to people with love and compassion.

I would like to thank all of my teachers in the Cognitive Science department, who guided me correctly and wisely on a subject I have just started to learn. Especially, I would like to express my gratitude to Prof. Dr. Hüseyin Cem Bozsahin and Asst. Prof. Dr. Murat Perit Cakır, for their valuable comments and suggestions during the thesis monitoring committee meetings. I would also like to thank Assoc. Prof. Dr. Cengiz Acartürk and Dr. Ceyhan Temürcü, who broadened my horizons with the lessons they gave.

I owe special thanks to my friends in Image Processing and Pattern Recognition Laboratory and lecturers in the Computer Engineering Department for their support and closeness throughout my studies

I would like to thank Prof. Dr. Sharlene Newman, from Indiana University, for providing us the TOL fMRI dataset and the preliminary information on complex problem solving.

I believe that the lessons I have learned from many people and many experiences in my 30 years of business life have also contributed to this work. The critical, guiding, and motivating interactions I had with my managers and colleagues, and the trainings I received and gave have also helped me to shape my way of thinking and approach towards issues and problems. I would like to especially thank my former manager Nuray Akmeriç, who had the biggest share in this.

I would like to express my gratitude to my parents who instilled in me the values of striving to be a good, productive and hardworking person. They taught me that honesty and integrity are what matters the most, and success does not count unless it is earned fair and square. Finally, I want to thank my husband Kemal Degirmendereli who unconditionally supported me during my journey and my daughter Deniz Degirmendereli who always brings joy and enriches my life.

TABLE OF CONTENTS

ABSTRACT iv

ÖZ vi

DEDICATION viii

ACKNOWLEDGMENTS ix

TABLE OF CONTENTS x

LIST OF TABLES xv

LIST OF FIGURES xvi

LIST OF ABBREVIATIONS xxi

CHAPTERS

1 INTRODUCTION 1

1.1 Motivation and Problem Definition 2

1.2 Measuring the Information Content of Neural Activities 4

1.3 Conceptual and Methodological Problems 6

1.3.1 Brain Complexity 6

1.3.2 fMRI Data Analysis Challenges 7

1.3.3 Statistical Stability and Data Sufficiency of Experiments 8

1.3.4 Conceptual and Methodological Gaps 8

1.4 Contributions of the Thesis 9

1.5 Organization of the Thesis 10

2 HUMAN BRAIN AS AN INFORMATION GENERATOR AND PROCESSOR 13

2.1 How Does Brain Code Information? 14

2.1.1 Main Elements of Neural Communication 14

2.1.2 Neural Codes 15

2.2 Measuring the Brain Activity 16

2.2.1 Neuroimaging Techniques 17

2.2.1.1 Functional Magnetic Resonance Imaging (fMRI) 18

2.2.1.2 Electroencephalography (EEG) and Magnetoencephalography (MEG) 20

2.2.1.3 Functional Near-Infrared Spectroscopy (fNIRS) 20

2.2.1.4 Positron Emission Tomography (PET) 21

2.3 Decoding Functional Brain Imaging Data 21

2.3.1 fMRI Data Analysis Process 22

2.3.2 Univariate and Multivariate Analysis Approaches 24

2.3.3 Major fMRI Data Analysis Methods 25

2.3.3.1 Model-driven Methods 26

2.3.3.2 Data-Driven methods 27

2.3.3.2.1 Decomposition-Based Methods 27

2.3.3.2.2 Clustering Analysis Methods 28

2.3.3.2.3 Graph Based Analysis 29

2.3.3.2.4 Artificial Neural Networks 31

2.3.3.2.5 Information Theoretic Methods 32

2.4 Chapter Summary 33

3 AN INFORMATION THEORETIC APPROACH FOR MODELING COGNITIVE STATES USING FMRI IMAGES 35

3.1 Measuring the Information by Shannon Information Theory 35

3.1.1 Shannon Model of Communication 36

3.1.2 Entropy in Thermodynamics and Information Theory 37

3.1.3 Entropy as an Average Information 38

3.1.4 Estimation of Information Content 38

3.1.4.1 Self-Information (Shannon Information) 39

3.1.4.2 Entropy Estimation as an Average Information 40

3.2 Information-Theoretic Interpretation of Neural Activity 41

3.2.1 Information-Theoretic Methods for fMRI Analysis 42

3.2.2 Information-Theoretic Applications in Neuroimaging Data Analysis 43

3.3 Chapter Summary 47

4 A NEW INFORMATION THEORETIC METHOD FOR REPRESENTING MENTAL STATES OF COMPLEX PROBLEM SOLVING 49

4.1 Complex Problem Solving Process 50

4.2 TOL Game as a Complex Problem Solving Experiment 50

4.3 Information Content of Brain Regions 52

4.3.1 Estimating Static and Dynamic Entropy for Brain Regions 53

4.3.1.1 Dynamic Entropy Estimation for Brain Regions Using Voxel Time Series 56

4.3.1.2 Static Entropy Estimation for Brain Regions 58

4.3.1.2.1 Static Entropy Estimation Using Representative Time Series 58

4.3.1.2.2 Static Entropy Estimation Using Voxel Time Series 61

4.4 Estimating Static and Dynamic Brain Networks 63

4.4.1 Kullback-Leibler Divergence Estimation 65

4.4.1.1 Estimation of Static Kullback-Leibler Divergence 66

4.4.1.2 Estimation of Dynamic Kullback-Leibler Divergence . . 67

4.5 Chapter Summary 69

5 COMPLEX PROBLEM SOLVING EXPERIMENT AND DATA ANALYSIS 71

5.1 TOL Experimental Setup and Nature of fMRI data 71

5.2 Entropy of Brain Anatomical Regions 73

5.2.1 Dynamic Entropy Variations of Complex Problem Solving and Resting States 73

5.2.2 Dynamic Entropy Variations of Resting State, Planning and Execution Subtasks for Expert and Novice Players 75

5.2.2.1 Comparison of BOLD and Dynamic Entropy Variations of Resting State, Planning and Execution Subtasks 75

5.2.3 Analysis of Low Entropy Regions in Complex Problem Solving 76

5.2.4 Static Entropy Analysis for Planning and Execution Subtasks . 79

5.2.5 Static Entropy Analysis for Expert and Novice Subjects 82

5.2.5.1 Expert-Novice Differences in Problem-Solving Skills . . 82

5.2.5.2 Active Brain Regions for Experts and Novices During TOL Problem-Solving 87

5.3 Static and Dynamic Brain Networks Analysis 90

5.3.1 Prominent Connections for Successful Sessions during the Planning and Execution Phases 92

5.3.2 Planning and Execution Phases Prominent Connections for Unsuccessful Sessions 92

5.3.3 Analysis of Brain Network Properties 95

5.3.3.1 Node Degree 95

5.3.3.2 Closeness centrality 95

5.3.3.3 Betweenness centrality 98

5.3.3.4 Common Prominent Regions According to Brain Network Properties 99

5.4 Classification of Planning and Execution tasks of Complex Problem Solving 100

5.4.1 Planning and Execution Tasks Classification Using Entropy and BOLD Values 100

5.4.2 Planning and Execution Tasks Classification Using Kullback-Leibler and Pearson Correlation Matrices 102

5.5 Summary and Interpretation of Results 103

6 CONCLUSION AND FUTURE DIRECTIONS 107

6.1 Critiques of the Proposed Method 108

6.2 Future Work 110

REFERENCES 115

APPENDICES

LIST OF TABLES

Table 5.1 Comparison of durations during planning and execution phases for successful and unsuccessful sessions. 86

Table 5.2 Classification performances of SVM for planning and execution tasks 101

Table 5.3 Within-Subject Performances for Planning and Execution Tasks 103

LIST OF FIGURES

Figure 2.1 The main elements of neural communication. Postsynaptic currents move through dendrites towards the cell body. If the current exceeds a threshold at the axon hillock, a spike is generated. It moves along the axon until it causes the release of neurotransmitters at the end of the axon. Neurotransmitters bind to matching receptors in the dendrites of connected neurons, causing ion channels to open. These channels allow the creation of postsynaptic currents in the receiving cell, and the process continues [1]. 15

Figure 2.2 A spatial and temporal comparison of the functional neuroimaging modalities [2]. 18

Figure 2.3 Digitally enhanced MRI images of the human brain [34]. 19

Figure 2.4 The main advantages and disadvantages of neurocognitive brain imaging techniques that are most commonly used in neuroscience research [3]. 22

Figure 2.5 The fMRI data processing pipeline that illustrates the different steps involved in a standard fMRI experiment [62]. 23

Figure 2.6 An array representation of voxel-specific MR signal time-series of a single participant and a single fMRI experimental run. Rows represent data acquisition times and columns represent voxels, i.e., three-dimensional locations. n_T and n_V denote the total number of time-series and total number of voxels, respectively. 24

Figure 2.7 fMRI data analysis with brain graphs. (a) First, the imaging data is preprocessed. Then, the brain is partitioned into regions, mostly using a brain atlas. Each region is assigned a regional representative time series. (b) A labeled simple graph is generated from the regional time series, where edge labels correspond to statistical dependency between brain regions, and brain regions are mapped to graph vertices. (c) The graph is embedded into a vector space. (d) statistical machine learning can be used for decoding the cognitive states. (e) Graph statistics can be estimated to analyze the brain graph [4]]. 30

Figure 3.1 The information-theoretic model of a communication system [5]. 37

Figure 3.2 Entropy plot for two possibilities with probabilities $p(x_i)$ and (1 – $p(x_i)$). 41

Figure 4.1 A set of Tower of London (TOL) problems that can be solved in three to six moves [6]. 51

Figure 4.2 The histogram (left) and the corresponding kernel probability density estimate (right) graphs plotted based on the same data values. Six kernels corresponding to six data points are indicated by red dashes. The blue line shows the resulting kernel density curve [7]. 55

Figure 4.3 a) Kernel density estimation with the *optimal bandwidth* (left), b) a *small bandwidth* (middle), and c) a *large bandwidth* (right) [8]. 56

Figure 4.4 Probability density functions for Right Middle Frontal gyrus for subject-226 session-1 (Left), and Right Caudate for subject-146 session-3 (Right). The pdf's are plotted using the planning time series in the session. The pdf's for each planning time instant are shown on the same graph. 58

Figure 4.5 The algorithm for the estimation of dynamic entropy for an anatomical region in a time-instance. 59

Figure 4.6 Probability density functions using two different bandwidths when the representative time series are used as random variable, for *Left Medial Orbitofrontal Cortex* across all subjects (Left) versus single subject (Right). The red dashes corresponds to a small bandwidth, blue line corresponds to a large bandwidth. In each plot Gaussian Kernels are used. Note that across subject density is a Gaussian mixture whereas the single subject density is a unimodal Gaussian function. . . . 61

Figure 4.7 The algorithm that summarizes the processes for the estimation of static entropy of the anatomical regions using representative time series for the planning task of a subject. 62

Figure 4.8 The algorithm that summarizes the processes for the estimation of static entropies of the anatomical regions by task, using voxel time series for the successful runs. 64

Figure 4.9 The algorithm for estimation of Kullback-Leibler divergences to generate the directed static brain networks for the planning task of a subject. 68

Figure 4.10 The algorithm for estimation of Kullback-Leibler divergences to generate arc-weights for the directed and undirected dynamic brain networks. 70

Figure 5.1 Averages of BOLD values across all sessions, based on anatomical regions with a 95% confidence interval. 72

Figure 5.2 Dynamic Entropy vs. time for a minimum entropy region: Entropy fluctuations (blue) for **Left Superior Parietal** during TOL puzzle solving (averaged for 12 subjects, 42 sessions). *Play* shows the time instances when the subject plays the TOL game and *Rs* shows, when the subject is back to the resting state, in red plot. The first puzzle begins at 33. time-instance 74

Figure 5.3 Dynamic Entropy vs. time for a maximum entropy region: Entropy fluctuations (blue) for **Right Insula** during a TOL session (averaged for 12 subjects, 42 sessions). *Play* shows the time instances when the subject plays the TOL game and *Rs* shows, when the subject is back to the resting state, in red plot. 74

Figure 5.4 Expert Player: Dynamic Entropy fluctuations of a low entropy region, Right Precuneus during session-3 of an expert player, Subject-175 (blue). *Pl* shows the time instances when the subject is doing Planning, *Ex* shows the execution and *Rs* shows, when the subject is back to the resting state, in the red plot. 76

Figure 5.5 Novice player: Dynamic Entropy fluctuations for a low entropy region, Right Precuneus during first session for a novice player, Subject-881 (blue). *Pl* shows the time instances when the subject is doing Planning, *Ex* shows the execution and *Rs* shows, when the subject is back to the resting state, in the red plot. 76

Figure 5.6 Expert Player: Dynamic Entropy fluctuations of a low entropy region, Right Lingual during session-3 of an expert player, Subject-175 (blue). *Pl* shows the time instances when the subject is doing Planning, *Ex* shows the execution and *Rs* shows, when the subject is back to the resting state, in the red plot. 77

Figure 5.7 Novice player: Dynamic Entropy fluctuations for a low entropy region, Right Lingual during first session for a novice player, Subject-881 (blue). *Pl* shows the time instances when the subject is doing Planning, *Ex* shows the execution and *Rs* shows, when the subject is back to the resting state, in the red plot. 77

Figure 5.8 BOLD time series and CPS task correlation for Subj-146 sess4 (Expert player) for an activated region Right Superior Parietal during TOL problem-solving. *Pl* shows the time instances when the subject is doing Planning, *Ex* shows the execution and *Rs* shows, when the subject is back to the resting state, in the red plot 78

Figure 5.9 Entropy values and CPS task correlation for Subj-146 sess4 (Expert player) for an activated region Right Superior Parietal during TOL problem-solving. *Pl* shows the time instances when the subject is doing Planning, *Ex* shows the execution and *Rs* shows, when the subject is back to the resting state, in the red plot 78

Figure 5.10 The lowest static entropy regions for planning and execution subtasks based on all voxels. The anatomical regions are sorted with respect to the lowest planning phase entropy values. Static Entropy estimation was made using the region representative time series averaged over all the voxels which resides in an anatomical region. 81

Figure 5.11 The lowest static entropy regions for planning and execution subtasks based on most informative voxels, selected by ANOVA method. The anatomical regions are sorted with respect to the lowest planning phase entropy values. Static Entropy estimation was made using the region representative time series by averaging only the most informative voxels, selected by ANOVA. 81

Figure 5.12 The lowest static entropy regions for planning and execution subtasks for successful runs based on most informative voxels, selected by ANOVA. The anatomical regions are sorted with respect to lowest planning phase static entropy values. Static Entropy estimation was made using the region representative time series,by averaging only the most informative voxels, selected by ANOVA. 83

Figure 5.13 The lowest static entropy regions for planning and execution tasks for unsuccessful runs based on most informative voxels (sorted by minimum planning entropy for successful runs). 83

Figure 5.14 Visualization of anatomical regions with the lowest static entropy measures for planning and execution phases based on most informative voxels, selected by ANOVA. Hot colors represent relatively low static entropy regions. The size of the circles are inversely proportional to the static entropy measures 84

Figure 5.15 The anatomical regions with the lowest static entropy values of expert and novice subjects based on most informative voxels, selected by ANOVA method. The anatomical regions are sorted with respect to the lowest planning phase entropy values. 88

Figure 5.16 Visualization of anatomical regions with the lowest static entropy measures for expert and novice players based on most informative voxels, selected by ANOVA. Hot colors represent relatively low static entropy regions. The size of the circles are inversely proportional to the static entropy measures. 89

Figure 5.17 Low entropy regions for expert and novice subjects for **planning phase** of TOL game. 89

Figure 5.18 Low entrpy regions for expert and novice subjects for **execution phase** of TOL game. 90

Figure 5.19 Static K-L divergences lower than a specific threshold (0.025) value between anatomical regions for *planning* phase (Left) versus the *execution* phase (Right) for subject-209. 91

Figure 5.20 Static K-L divergences (Left) and Pearson Correlation matrix (Right) during a specific time instance for Subject-209 session-3. 92

Figure 5.21 Undirected connection graph which shows most prominent connections for **planning** phase for **successful** TOL problem-solving sessions. 93

Figure 5.22 Undirected connection graph which shows most prominent connections for **execution** phase for **successful** TOL problem-solving sessions. 93

Figure 5.23 Visualization of anatomical regions which have shortest K-L distance measures for **planning** phase for **successful** sessions. 94

Figure 5.24 Visualization of anatomical regions which have shortest K-L distance measures for **execution** phase for **successful** sessions. 94

Figure 5.25 Undirected connection graph that shows most prominent connections for **planning** phase for **unsuccessful** TOL problem-solving sessions. 96

Figure 5.26 Undirected connection graph that shows most prominent connections for **execution** phase for **unsuccessful** TOL problem-solving sessions. 96

Figure 5.27 Visualization of anatomical regions which have shortest K-L distance measures for **planning** phase for **unsuccessful** sessions. 97

Figure 5.28 Visualization of anatomical regions which have shortest K-L distance measures for **execution** phase for **unsuccessful** sessions. 97

Figure 5.29 Number of connections with other regions to a region during planning and execution tasks. 98

Figure 5.30 The regions that have high closeness centrality for planning and execution tasks. 99

Figure 5.31 Nodes with high betweenness centrality for planning and execution tasks. 100

Figure 5.32 Classification performances (%) of SVM classifier for planning and execution tasks. 101

Figure 5.33 Within-subject classification performance (%) of SVM classifier for planning and execution tasks. 102

LIST OF ABBREVIATIONS

CPS	Complex Problem Solving
TOL	Tower of London
fMRI	Functional Magnetic Resonance Imaging
ER-fMRI	Event-Related fMRI
rs-fMRI	Resting state fMRI
EEG	Electroencenography
MEG	Magnetoencephalography
fNIRS	Functional Near-Infrared Spectroscopy
PET	Positron Emission Tomography
SPECT	Single-Photon Emission Computed Tomography
AAL	Automated Anatomical Labeling
ERP	Event Related Potential
MVPA	Multi-Voxel Pattern Analysis
ANOVA	Analysis of Variance
SVM	Support Vector Machines
KNN	K-Nearest Neighbors
SNR	Signal-to-Noise Ratio
HRF	Hemodynamic Response Function
BOLD	Blood Oxygen-Level Dependent
TR	Repetition Time
GLM	General Linear Model
ICA	Independent Component Analysis

PCA	Principal Component Analysis
ApEn	Approximate Entropy
pdf	Probability density function
ROI	Region of Interest
SPM	Statistical Parametric Mapping
MI	Mutual Information
BCI	Brain-Computer Interface
MSE	Multiscale Entropy
MPE	Multiscale Permutation Entropy
SampEn	Sample Entropy
ApEn	Approximate Entropy
BENtbx	Brain Entropy Mapping Toolbox
KDE	Kernel Density Estimation
IQR	Interquartile Range
K-L	Kullback-Leibler (Divergence)
PreCG.L	Left Precental gyrus
PreCG.R	Right Precental gyrus
SFGdor.L	Left Superior frontal gyrus, dorsolateral
SFGdor.R	Right Superior frontal gyrus, dorsolateral
ORBsup.L	Left Superior frontal gyrus, orbital part
ORBsup.R	Right Superior frontal gyrus, orbital part
MFG.L	Left Middle frontal gyrus
MFG.R	Right Middle frontal gyrus
ORBmid.L	Left Middle frontal gyrus, orbital part
ORBmid.R	Right Middle frontal gyrus, orbital part
IFGoperc.L	Left Inferior frontal gyrus, opercular part

IFGoperc.R	Right Inferior frontal gyrus, opercular part
IFGtriang.L	Left Inferior frontal gyrus, triangular part
IFGtriang.R	Right Inferior frontal gyrus, triangular part
ORBinf.L	Left Inferior frontal gyrus, orbital part
ORBinf.R	Right Inferior frontal gyrus, orbital part
ROL.L	Left Rolandic operculum
ROL.R	Right Rolandic operculum
SMA.L	Left Supplementary motor area
SMA.R	Right Supplementary motor area
OLF.L	Left Olfactory cortex
OLF.R	Right Olfactory cortex
SFGmed.L	Left Superior frontal gyrus, medial
SFGmed.R	Right Superior frontal gyrus, medial
ORBsupmed.L	Left Superior frontal gyrus, medial orbital
ORBsupmed.R	Right Superior frontal gyrus, medial orbital
REC.L	Left Gyrus rectus
REC.R	Right Gyrus rectus
INS.L	Left Insula
INS.R	Right Insula
ACG.L	Left Anterior cingulate and paracingulate gyri
ACG.R	Right Anterior cingulate and paracingulate gyri
DCG.L	Left Median cingulate and paracingulate gyri
DCG.R	Right Median cingulate and paracingulate gyri
PCG.L	Left Posterior cingulate gyrus
PCG.R	Right Posterior cingulate gyrus
HIP.L	Left Hippocampus

HIP.R	Right Hippocampus
PHG.L	Left Parahippocampal gyrus
PHG.R	Right Parahippocampal gyrus
AMYG.L	Left Amygdala
AMYG.R	Right Amygdala
CAL.L	Left Calcarine
CAL.R	Right Calcarine
CUN.L	Left Cuneus
CUN.R	Right Cuneus
LING.L	Left Lingual gyrus
LING.R	Right Lingual gyrus
SOG.L	Left Superior occipital gyrus
SOG.R	Right Superior occipital gyrus
MOG.L	Left Middle occipital gyrus
MOG.R	Right Middle occipital gyrus
IOG.L	Left Inferior occipital gyrus
IOG.R	Right Inferior occipital gyrus
FFG.L	Left Fusiform gyrus
FFG.R	Right Fusiform gyrus
PoCG.L	Left Postcentral gyrus
PoCG.R	Right Postcentral gyrus
SPG.L	Left Superior parietal gyrus
SPG.R	Right Superior parietal gyrus
IPL.L	Left Inferior parietal gyri
IPL.R	Right Inferior parietal gyri
SMG.L	Left Supramarginal gyrus

SMG.R	Right Supramarginal gyrus
ANG.L	Left Angular gyrus
ANG.R	Right Angular gyrus
PCUN.L	Left Precuneus
PCUN.R	Right Precuneus
PCL.L	Left Paracentral lobule
PCL.R	Right Paracentral lobule
CAU.L	Left Caudate nucleus
CAU.R	Right Caudate nucleus
PUT.L	Left Lenticular nucleus, putamen
PUT.R	Right Lenticular nucleus, putamen
PAL.L	Left Lenticular nucleus, pallidum
PAL.R	Right Lenticular nucleus, pallidum
THA.L	Left Thalamus
THA.R	Right Thalamus
HES.L	Left Heschl gyrus
HES.R	Right Heschl gyrus
STG.L	Left Superior temporal gyrus
STG.R	Right Superior temporal gyrus
TPOsup.L	Left Temporal pole: superior temporal gyrus
TPOsup.R	Right Temporal pole: superior temporal gyrus
MTG.L	Left Middle temporal gyrus
MTG.R	Right Middle temporal gyrus
TPOmid.L	Left Temporal pole: middle temporal gyrus
TPOmid.R	Right Temporal pole: middle temporal gyrus
ITG.L	Left Inferior temporal gyrus
ITG.R	Right Inferior temporal gyrus

CHAPTER 1

INTRODUCTION

> The noblest pleasure is the joy of understanding.
>
> Leonardo da Vinci

Since the beginning of mankind, human's greatest quest is to understand themselves. When the Greek philosopher Socrates was asked to summarize what all philosophical decrees could be reduced to, he replied "Know thyself", referring to the maxim inscribed in the Temple of Apollo in Delphi. The word "Know yourself", which accentuates that the most important pursuit of our lives is to discover who we really are, has been expressed by many thinkers and philosophers throughout history.

A Turkish folk poet and a Sufi Dervish Yunus Emre said, "Knowledge is knowing yourself. If you do not know yourself, what is the use of this knowledge?". Another Turkish poet, a Sufi, Haji Bayram Wali (Haci Bayram Veli) has an impressive poem called, "You know yourself". The Austrian writer Stefan Zweig said that "once a man has understood the humanity in himself, he will understand all human beings". All this rhetoric emphasizes how crucial it is for the human to know himself. It is apparent that the way to understand oneself is to understand one's own mind.

The desire to understand our own minds has been the driving force behind many scientific efforts. For centuries, scientists and philosophers have tried to explain the enigma of the human mind, however, until the last century, the only basis for understanding and interpreting the mind was human discourse and behavior. When it became clear that various cognitive disorders originate in the brain, it was realized that mental phenomena have a neurophysiological basis and that the brain is an important tool for understanding the mind.

Since the late 19th century, it has been accepted that this three-pound organ is the place of thought, intelligence, emotions, the interpreter of the senses and the controller of behavior, even the source of all the traits that define our shared humanity.

Until the mid 20th century, human brain was considered almost incomprehensible. However, due to the rapid advances in neurological, cognitive and behavioral science and the development of new research techniques, more information about the brain has been learned especially in the last 30 years than in any previous period. The

improvement of neuroimaging technologies has allowed the study of the brain to go beyond just behavioral experiments, making it possible to observe what is happening inside the brain. This approach led to the study of cognition, intelligence and behaviors, based on how the nervous system represents and processes information. It also offers an implementation base to confirm independently developed cognitive and psychological theories.

Although deciphering the brain's activity using neuroimaging techniques is considered a giant step towards understanding the mind, the link between the brain and mind is still a fundamental question that has not been fully resolved. So far, thousands of studies have examined how the human brain, and therefore the mind, works from different perspectives, but we are still far from fully understanding it.

Nowadays, we can record the electrochemical activities of millions of neurons in the brain while it's operating, and we can collect large volumes of neural data. Additionally, various sophisticated data analysis techniques have been developed to extract "meaningful" information from this data to reveal how mental representations match patterns of neural activity. These techniques qualitatively and quantitatively evaluate the activated brain regions, and the trajectory of information flow within the brain. However, we still do not have a comprehensive method which utilizes neural data to the full extent to explain precisely how information generated, represented and processed in the human brain. We lack models of interaction among different brain regions for various cognitive tasks such as planning, learning, reasoning, problem solving and decision making.

The key challenge is to understand how the electrochemical processes in our nervous system turn into our thoughts, behaviors, and perceptions of the physical world, we live in. A formal answer to this challenge requires a set of very complicated mappings of physical activities of the brain to mental activities. In other words, we need formal analysis techniques, models and conceptual frameworks, which can be tested and refined to manifest "correlations" between mental and neuronal processes.

A large amount of neural data enriches our descriptive information, but it does not necessarily mean high explanatory power, which requires principles with high heuristic potential [9]. Also, the techniques alone are short to interpret the data appropriately without the underlying theoretical context. Methods that are not aligned with the theoretical understanding would not be sufficient for a generalized explanation of the mental functions.

1.1 Motivation and Problem Definition

In this thesis, we aim to develop a computational model with an underlying theoretical context for localizing the brain regions that contribute to the execution of a high-level cognitive task, called complex problem solving (CPS). Then, we construct dynamic and static brain network representations among these regions to represent

their coordinated activities. We also aim to reveal and represent the cognitive states of the mind performing a complex problem solving task.

According to Simon & Newell (1971), "the term problem solving refers to a higher-order cognitive function directed toward identifying problems with the current state and generating and implementing potential solutions to achieve a goal state" [10]. Problem solving is not only necessary for the advancement of technology, but also for human survival. Therefore, it is important to understand their neural and cognitive bases. However, as a higher-level mental task, it requires the distributed activity of many parts of the brain. It interacts with many other cognitive processes such as imagining, searching, learning, judging, decision making, inference, analysis, and synthesis [11]. Therefore, the underlying mental events are difficult to localize and hard to measure properly. The development of a computational model based on a valid theory is needed to explain the mental mechanisms involved in complex problem solving.

In this thesis, we analyze the relationship between the neural activity of the brain and complex problem solving task using functional magnetic resonance imaging (fMRI) data, which allows us to noninvasively measure the brain activations, spatially (area of the activation), and temporally (time of the activation). This method, which was introduced for the human brain in 1990, is the most widely used brain imaging technique to investigate the living human brain while subjects perform tasks and experience mental states.

We examine the information content of the neural variability reflected in fMRI signals during the complex problem solving by the activities produced in anatomical regions. We assume that fMRI data provides a basis for studying the correlation between neural and cognitive processes. They allow us to use neurophysiology to understand the structure of mental processes by detecting which mental processes involve similar and different neurophysiological processes.

Human brain is described as a complex system composed of relatively specialized and domain-general components that operate in synchrony and coordination, performing different cognitive processes, and exchanging information in a hierarchical and integrated way. Based on this description, we investigate the question of whether cognitive processes can be isolated and defined in a way that allows them to be associated with certain brain regions. The findings that a particular region is selectively involved in a specific mental process are informative not only because it shows us where this process takes place, but also because it indicates that the brain has specialized mechanisms for a specific mental process.

Based upon the above analysis, we suggest a computational model, which estimates the active brain regions during a complex problem task. The suggested computational model is applied to an fMRI dataset recorded while subjects play the Tower of London (TOL) game. The details of this game is provided in Chapter 2.

As a second computational model, we suggest static and dynamic brain networks to estimate the interactions among the activated anatomical regions from the fMRI

data, recorded during the phases of complex problem solving. The estimated brain networks can successfully discriminate the phases of complex problem solving task by simple machine learning methods.

The computational models, suggested in this thesis enable us to observe the planning and execution phases of complex problem solving, from the fMRI data, as hypothesized by Simon and Newell, in [10]. They also provide a basis for understanding common traits and differences between individuals, when solving complex problems and allow distinguishing the differences in the brain activity of strong and weak problem solvers.

Our investigation may lead to a better characterization of the processes that support complex problem solving compared to classical statistical statistical techniques. It is expected that understanding the neural bases of a cognitive task, such as complex problem solving, provides valuable new insights into the mechanisms of internal knowledge representation and processing behind the mental processes. The development of computational models for analyzing the mental mechanisms involved in problem solving would greatly facilitate the explanation of this phenomenon.

Although our models are developed for investigating the complex problem solving task, based on fMRI data, it can be easily applicable to other cognitive tasks.

1.2 Measuring the Information Content of Neural Activities

The major assumption of this thesis is that the human brain is a system of integrated neural information processing units, where a group of neurons or brain regions works in an orchestral harmony. These activities produce and convey a great amount of information, which can be quantified and characterized by some information theoretic measures such as Shannon entropy and relative entropy.

The proposed model represents the brain activities by measuring the information content of neural signals. For this purpose, we employ the major concepts, such as entropy and relative entropy, defined in Shannon Information Theory. We employ the concepts to investigate the information content of neural activities in anatomical regions and that of the interconnections among them, based on fMRI data recorded while the subjects perform a complex problem solving task.

Shannon entropy is one of the most popular mathematical descriptions of how to quantify, represent, and communicate information effectively. It is a mathematically defined measure of the average amount of information required to represent an event.

Entropy has been recently seen as a promising tool for the analysis of the dynamic nature of neural signals [12]. It characterizes the voxel time series to aid the identification and quantification of regular and random signals. It helps to estimate the regularity and predictability of neural operations.

The basic questions in neural information processing is as follows:

- How do firing patterns of specific neurons and neural ensembles represent the external or internal stimuli?
- What do responses of neurons tell us about a stimulus?

In this study, we investigate related questions:

- How much information is conveyed during a cognitive process?
- How much does the response of neurons tell us about the outcome of a stimulus?

Shannon entropy enables us to find quantitative answers to the above questions.

Shannon entropy quantifies the amount of information required to specify the state of a system, or the degree of lack of information about the exact state of the system in question. In other words, it measures the amount of information that can potentially be gained (or the potential reduction in our uncertainty) once we have learned the outcome of a system.

The entropy of a dynamic system is expected to be high, if it has many possible states with equal likely probabilities. Therefore, a high entropy value indicates greater randomness or ambiguity of the system's state or outcome, while low entropy implies high regularity and predictability. Thus, zero entropy implies complete availability of the information about the system's state or outcome. We suggest that measuring the entropy of a brain region based on fMRI recordings provides us information about the mental states of that region.

In this study, we measure the information content of brain anatomical regions during two main phases of complex problem solving, namely, planning and execution, for expert and novice problem solvers using first-order Shannon entropy estimates from fMRI data.

Our major assumption is that the low entropy brain regions are more intimately involved in complex problem solving processes compared to the high entropy regions. Therefore, we examine the relationship between the phases of the problem solving task and the entropy measures of anatomical regions.

Our study proposes two new entropy definitions: static entropy and dynamic entropy. Static entropy is defined over an interval of time, while dynamic entropy measures the activity of a brain volume at each time instant. We also estimate the interactions among the anatomical regions using the relative entropy which enables us to extract dynamic brain networks.

Experimental neuroscience reveals that a collection of anatomical regions exhibit coordinated behavior to form a cognitive task. When we measure the entropy of an anatomical region, we expect that the voxel signal intensities that vary with a cognitive task in an organized manner, produce low entropy values. On the other hand, an

anatomical region whose voxel intensities vary in a relatively random pattern would not participate in the task, yielding high entropy values. In other words, we expect that low entropy indicates regular and organized behaviors of the underlying region, while high entropy regions exhibit relatively random behaviors. In this study, we investigate the relationship between low entropy measures and the activation in regions involved in complex problem solving.

We propose a novel brain network estimation method using Kullback-Leibler Divergence, which is also called relative entropy. The estimated static and dynamic brain networks model the interactions during the phases of the complex problem solving process. We test the validity of the suggested brain networks by training a classifier with the arc-weights of the estimated dynamic brain networks, and measuring the test performances of the planning and execution phases.

We intend to show that the computational model suggested in this study, can be used for the analysis of neural data, where entropy and relative entropy provides a powerful explanatory tool for investigating the underlying cognitive processes. We hope that the information theoretical concepts and methods proposed in this study provide a rigorous framework for analyzing many aspects of cognition and uncovering some of the underlying principles of information processing in the brain.

1.3 Conceptual and Methodological Problems

The development of neuroimaging techniques has led researchers to study how large-scale patterns of brain activity map to specific mental states or processes. This has already been partially accomplished by advanced technologies and methodologies that have allowed some degree of understanding of the brain-mind relationship. However, there are still many conceptual and methodological issues to be investigated in brain research and fMRI analysis.

1.3.1 Brain Complexity

Human brain can be considered as extremely complex information processing system, with a very high computing power. There are approximately 150-200 billion cells in our brain, which is 20 times more than the number of people on Earth. The folds in the brain's external surface increase the surface area and the number of possible connections between neurons which increases the brain's computing power.

In the brain, information is distributed across different abstraction levels and different cell types, from neurons to networks. Signals spreading out from various types of cells encode a large variety of information. There are several thousands of specific types of neurons across the anatomic regions, each of which has its own unique role. We are far from explaining the functionality of all types of neurons and other brain cells. There is no well-defined categorization of these neurons, unlike the periodic tables of chemistry.

Furthermore, neurons with similar characteristics can produce different actions based on how they connect with each other. It is reported that there are hundreds of different kinds of neurotransmitters and many different kinds of receptors that mediate the neuronal connections [13]. Various combinations of the neurotransmitters and receptors can lead to different types of connectivity, which forms the connectome of the human brain. An electrochemical signal transmitted by a neuron may have different effects on other neurons to which it is linked, depending on the mediating neurotransmitters and receptors. Thus, it can encode different information in each case. These issues make it very difficult to unravel neuronal connectivity and signaling.

While contemporary imaging technologies allow us to observe complex patterns of neuronal activity associated with specific cognitive or behavioral states, as more and more neuroscientific data emerges, it becomes critical to question the limitations and capabilities of neuroscientific data for analysis of human brain.

1.3.2 fMRI Data Analysis Challenges

fMRI technique measures oxygenation and flow changes in the blood in response to neural activity. When the neurons in a brain region are activated, they consume more oxygen. The blood flow in the active areas increases to meet this demand. The fMRI signals measure this response, which is called Blood Oxygen-Level Dependent (BOLD) response. Therefore, the BOLD response provides only an indirect measurement of neural activation. Although it is assumed that the BOLD response increases with neural activation, it is much slower than the neural activation, since neurons fire thousands of times faster than blood flows. As a result, the peak of the BOLD signal lags several seconds behind the peak neural activation. Thus, observing the brain activity during the rapidly changing cognitive states from the fMRI recordings is rather difficult.

Furthermore, fMRI measures are the average blood flow around a particular region of neurons, which is called a "voxel". Each voxel captures blood flow in the region of approximately 100.000 neurons and more than 4 million synapses over about one second. Therefore, fMRI maps should be viewed as a composite of all activity for a group of neurons occurring during a particular period of mental activity.

fMRI experiments produce a large amount of highly complex and noisy data due to dozens of experimental and measurement problems. Most fMRI data analysis techniques involve a separation method between 'signal' and 'noise' [9]. This separation method generally includes spatial and temporal averaging, filtering, smoothing, or normalization (see section 2.3.1) [14]. Noise reduction methods may result in loss of important information embedded in fMRI data. Correspondingly, the data is overlooked while extracting their 'useful' parts and ignoring the 'less important' parts. If we do not have an appropriate measure of 'usefulness', the approaches to focus only on the meaningful part of the brain signal may lead to the loss of crucial information already available.

In order to increase the generalizability of statistical analysis results, the fMRI brain imaging data is aggregated across multiple subjects. However, the brain structure, connections, and dynamics varies naturally among the people. A normalization procedure is used to map the brains into a standard brain template for making the fMRI data compatible across the subjects. This normalization may cause spatial imprecision and blurriness in group data, resulting in mislocalization of the activation in small structures. The development of advanced inter-subject normalization and improved smoothing techniques would help to avoid the negative effects of this process. However, in addition to variability among subjects, the variability of functional signals among experimental trials makes it difficult to compare the fMRI datasets properly.

1.3.3 Statistical Stability and Data Sufficiency of Experiments

It is known that collecting and analyzing neuroimaging data is very expensive. Therefore, in brain research, most experimental studies rely on a relatively small number of subjects and trials compared to the data obtained from experimental psychology for analysing the underlying cognitive process. For this reason, the results of many studies are not based on statistically stable and sufficient number of samples.

Due to the above mentioned problems, fMRI data analysis results may consist of many false positives leading to non-reproducible and sometimes misleading findings of the underlying cognitive phenomenon. A false positive may reveal the existence of a pattern, which can not be observed in another set of experiments.

An important factor that leads to the non-reproducibility problem is that new results are often seen as exciting and vital to the advancement of science. This is especially true in neuroscience, where substantial parts of the brain structure and function are still unexplored. Since scientific journals are focused on novelty rather than replication, the majority of scientific work disregards the test of reproduction of the same results.

1.3.4 Conceptual and Methodological Gaps

Brain studies involve a large variety of multidisciplinary research including neuroscience, cognitive science, psychology, computer science, medicine and biology. A scientific discipline can be characterized by its epistemic object such as mind and brain, concepts, theories, and methods [9]. Different disciplines also have different implicit assumptions. Neuroscientific studies are based on a very heterogeneous set of concepts, theories, models and methods. This interdisciplinary character makes conceptual, theoretical and methodological problems more visible [9]. Successful integration requires a regulation to interrelate different concepts, methodological strategies and contexts to be able to support each other.

Empirical research in neuroscience is mostly technology-driven. Methodology does not just mean listing the techniques used in research, but it is a theory of the epistemic

value of these techniques [9]. For example, statistical analyses only give quantitative descriptors of the relationships between the available values of the variables. The difficulty arises in assessing whether the analysis and results are meaningful for the relevant disciplines. The approach to explaining complex behaviors through the brain activity alone fails to take, for example, psychological or other factors into account and can mislead us.

Neither large volumes of data nor a wide range of advanced techniques can provide a complete understanding without establishing the general principles of the human brain and its functionalities. Establishing general principles is achieved by constituting a common terminology and an inclusive set of methodologies across the disciplines, which involve the brain research.

We need a holistic approach, which characterize the dynamics of the brain to understand the mind. Bridging theoretical, conceptual and methodological gaps among the related disciplines will be a big step toward this goal. The very major assumption of this thesis is that, information theory has the potential to both integrate and unify a wide range of phenomena in a single framework, while also facilitating specific hypothesis testing in a wide variety of brain research [15].

1.4 Contributions of the Thesis

In this thesis, we propose an information-theoretic approach for the representation of human brain dynamics to investigate the brain activities using fMRI data. The suggested approach offers computational models to measure the information content of the anatomical regions. It also offers models to estimate the dynamic brain networks at each time instance across the anatomical regions. Therefore, the suggested computational models enables a mapping between quantitative measures of the brain activity and the cognitive experience.

We can summarize the contributions of this thesis as follows:

- We develop a new computational model to localize the active brain regions, which contribute to the execution of complex problem solving task. The suggested computational model maps the neural activities measured by the fMRI recordings to the underlying cognitive processes of the subjects.
- We propose two measures to quantify the information content of anatomical regions: static entropy and dynamic entropy. These measures are adopted from Shannon entropy to estimate the activities related to the complex problem solving phases in anatomical regions.
- We propose a novel method using Kullback-Leibler Divergence for the representation of the static and dynamic brain networks, which estimates the strength of the relations among anatomical regions during complex problem solving. Traditional brain network research mostly employs various statistical correlations between brain areas to construct a brain network. On the other hand, our

approach measures the distance between the probability density functions (pdf) of anatomical regions to establish functional connections.

- In order to test the validity of the proposed computational models, we train a machine learning algorithm for the classification of planning and execution phases of complex problem solving. We report the classification performances above 90% in most experiments. These results can be considered as an indication of the validity of our computational models.
- The suggested method can be applicable to a wide range of event related fMRI data set. In this study, we suffice to apply our models to the fMRI data recorded during complex problem solving task.

1.5 Organization of the Thesis

In Chapter 2, we investigate the human brain as an information generation and processing system. Based on the basic structure of the nervous system, we examine the processing and transmission of neural information in the brain. We review popular neuroimaging techniques for measuring neural activity in the brain. Next, we describe common methods, used to decode and analyze neural signals recorded by neuroimaging tools to gain insights into cognitive processes corresponding to the measured neural activity. Finally, we discuss the pros and cons of current fMRI data analysis methods. This part provides the background and methodologies for the analysis of brain imaging data.

Chapter 3 provides the information theoretic tools used in this study based on fMRI data. First, we describe the Shannon communication model and interpret this model in the neuroscience context. We explain the information entropy concept, which is the fundamental measure of information theory. Then, we explore the connection of the human brain and mind with the concepts of entropy. We investigate how these concepts can be applied to the study of the dynamics of brain activity. In the last part, we propose a computational model for analyzing the fMRI data. We went through some examples of information-theoretic studies that have been carried out in neuroimaging data analysis.

In Chapter 4, we discuss our proposed computational model in detail. First, we briefly describe the complex problem solving process and its main phases, namely planning and execution. Then, we present our data-driven method that provides the information theoretic representation of the brain. We explain how we apply this method in the context of complex problem solving for exploring the neural and mental underpinnings of this process. This part describes the suggested entropy measures to identify activated brain regions and reveal brain states associated with a cognitive process. We define two types of entropy: static entropy and dynamic entropy for the analysis of the neural activation. We also propose a new method using Kullback-Leibler Divergence (relative entropy) for the estimation of the static and dynamic brain networks, during the planning and execution phases of the complex problem solving task.

In Chapter 5, we present the main findings of the data analysis for the proposed method described in the previous chapter. First, we explain the Tower of London (TOL) experimental setup, the fMRI data collection environment and the fMRI data properties. Then, we examine the behavior of static and dynamic entropy variations across subjects, the main phases of problem-solving, and the anatomical regions. We estimate the entropy of anatomical regions with low entropy. We suggest that the low entropy regions are activated by the complex problem solving task. We compare the low entropy anatomical regions for strong and weak problem solvers in order to understand neural activation discrepancies. Then, we discuss what is behind the differences in the ability to solve a complex problem, thereby, examine the differences between experts and novices in problem-solving skills. In the next part, we present the results of the static and dynamic brain networks analysis. We estimate planning and execution phases prominent connections for successful and unsuccessful runs, using the Kullback-Liebler Divergence method. We test the validity of the suggested brain networks by training an SVM classifier with the arc-weights and compare our method with the other network models.

In Chapter 6, we summarize our research findings and critique the proposed computational model. We also provide a future direction for research.

CHAPTER 2

HUMAN BRAIN AS AN INFORMATION GENERATOR AND PROCESSOR

> Brain, body, and world are united in a complex dance of circular causation and extended computational activity.
>
> Andy Clark

We perceive our environment with our sensory organs, generate vital information, determine our behavior, think about our actions, imagine, and plan the future in our brain. A system retrieves, processes and stores information, if it has a language for an internal representation, which encodes the information.

For more than fifty years, neuroscientists have been in intense debate about how information is encoded and transmitted in the human brain. While it is known that information is represented in our brain through the electrical activity of neurons, the details of this representation, called "neural coding", have not yet been fully deciphered [16]. How electrical pulses are transformed into emotions, thoughts, and ideas is still an important question to be answered. The development of neuroimaging techniques, and the techniques for simulating the brain processes by computational models play an important role to address this question.

This chapter attempts to roughly answer the following questions:

- How does the brain generates and transmits neural information?
- How do we measure neural activity in the brain?
- How do we extract information about the neural activities from the neural activity measurements using neuroimaging data analysis methods?

The first part of this chapter explains the structure of the brain cells. We overview the way cells generate and transmit signals. The next part describes the neuroimaging techniques for measuring the activity in the brain. There are various techniques used for capturing signals produced and transmitted by the nerve cells. This part overviews and compares the most frequently used functional neuroimaging techniques such as PET, SPECT, fMRI, fNIRS, EEG and MEG. Next, we explain the methods used to

analyze the signals recorded by neuroimaging tools. In this part, we provide the background and methodologies for the analysis of brain imaging data.

2.1 How Does Brain Code Information?

The human brain is suggested as a complex, physical, dynamic information processing system, that receives, processes, generates, and transmits information. In most of the information processing systems, the issues to consider are what is encoded, what is the code used to transmit the information, how noisy the code is, and how the information is decoded.

2.1.1 Main Elements of Neural Communication

Information processing in the brain occurs with the joint and coordinated activity of many neural elements. Information is distributed across many cell types, from neurons to networks, with contributions from various neural elements. The basis of information processing is the information received, transformed, and transmitted by brain cells. There are two types of cells in nervous system: *neurons* that are considered as the basic computing units of the brain, and *glial cells* that surround and support them. Neurons are tree-like structures with a small *cell body*, an *axon* which is a long, nerve fiber, and large branch-like *dendrites* as schematically shown in Figure 2.1. The axon carries signals to other neurons; dendrites receive input signals from other neurons or from sensory receptors.

The basic unit of communication and coding in the brain is the *spike* or *action potential*, generated by a neuron [17]. A spike is an electrical impulse of about a tenth of a volt that lasts for less than a millisecond [17]. The currents received by the dendrites move from the dendrites towards the cell body of the neuron. At the end of the cell body is the axon hillock, which controls the firing of the neuron. If the sum of the signal strengths exceeds a certain threshold, the neuron generates a typical spike (action potential). This electric current spreads rapidly along the axon. The arrival of this current causes a transmitter to be released at the axon terminal, which usually affects the membrane of the target neuron.

The specialized area where neurons communicate with each other is called a *synapse*. It is where a transmitter from a neuron is released. Neurotransmitters bind to matching receptors in the dendrites of connected neurons [1]. Thus, through synaptic transmission, the electrical signal in a neuron passes from the terminal of its axon to another cell. Besides the electrical signal, there are other signaling forms that stem from neurotransmitter diffusion. Synaptic connections between neurons allow information transfer by interconnecting neurons to form the circuitry The activity of a single neuron may be triggered by thousands of synaptic inputs.

Each neuron can be considered as a node of networks with one or more specific cognitive functions. Nerve cells with similar properties produce different actions de-

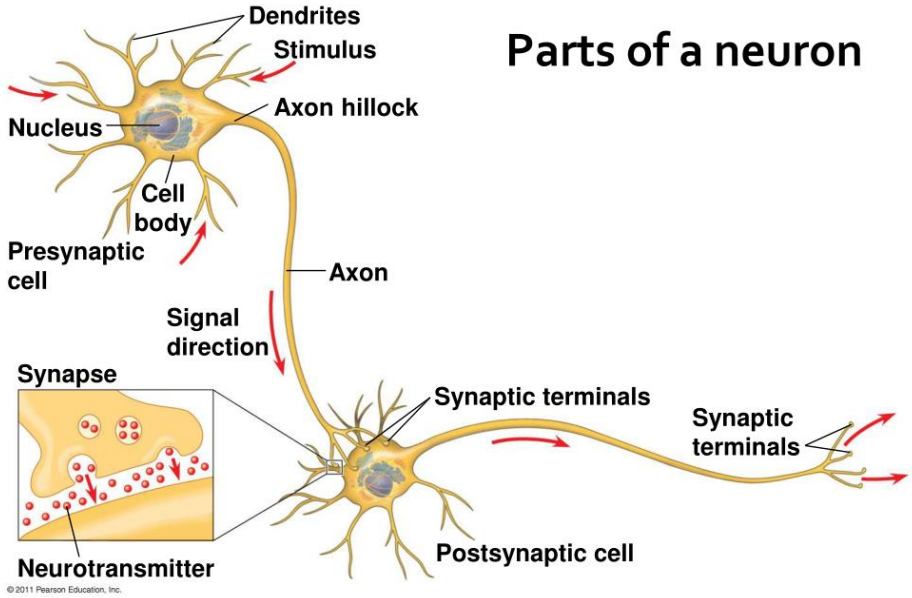


Figure 2.1: The main elements of neural communication. Postsynaptic currents move through dendrites towards the cell body. If the current exceeds a threshold at the axon hillock, a spike is generated. It moves along the axon until it causes the release of neurotransmitters at the end of the axon. Neurotransmitters bind to matching receptors in the dendrites of connected neurons, causing ion channels to open. These channels allow the creation of postsynaptic currents in the receiving cell, and the process continues [1].

pending on how they connect with each other. It has also been shown that there are as many as ten thousand specific types of neurons in the brain, each of which has its own specific role. The spikes emanating from different types of neurons are assumed to encode different types of information. Furthermore, there are many different types of neurotransmitters and receptors. Various combinations of these can lead to different types of currents. Therefore, a spike transmitted by an axon can be received by many different neurons, and depending on the mediating neurotransmitters and receptors, it can have different types of effects on each neuron, thus, it can code different information in each case [1].

Figure 2.1 summarizes the main elements of neuronal communication.

2.1.2 Neural Codes

The sequence of spikes generated by neurons carry information from one site to another in the brain. This train of spikes or action potentials can be considered as an

element of a *neural code* [18]. The action potential is an all-or-none event and is conducted without decrement [19]. It is supposed that each bit of information is encoded in the spatial and temporal patterns of these electrical impulses.

Rather than a single universal code that explains what patterns of spikes mean, neurons are supposed to generate many types of codes, depending on what sort of information to encode. However, there are some general principles. In the standard model, there are two kinds of neural codes: namely a *rate code* and a *temporal code*. In the rate code, information is encoded only by the average rate of firing. In the temporal code, information is also encoded in the precise timing of each pulse, with the sensitivity of the submillisecond range; in the temporal firing, the patterns with the same average rate can encode different messages [16]. The temporal resolution of the nerve code has been detected to be on a millisecond time scale, indicating that precise spike timing is an important element in neural coding [20, 21].

2.2 Measuring the Brain Activity

Observing the active areas of the human brain during cognitive processes is very crucial to reveal the biological underpinnings of cognitive functioning. The development of neuroimaging techniques made it possible to study the correlation between our brain and our cognition and had unprecedented insights into the complexities of the working brain. Invivo-brain imaging allows us to observe large-scale systems of interconnected brain regions involved in various cognitive processes such as speech, vision, hearing, learning and thinking.

The development of functional imaging techniques such as positron emission tomography (PET), single-photon emission computed tomography (SPECT), functional magnetic resonance imaging (fMRI), electroencephalography (EEG), and magnetoencephalography (MEG) has provided great opportunities for the study of a wide range of brain functions. The application of these techniques has led to significant advances in uncovering the way the brain process information while performing certain tasks such as perception, attention, memory, reasoning, or language.

Localization of the cognitive functions has a long history in brain-related studies. Early localization theorists included the neuroanatomist-physiologist Franz Gall (1758-1828) and phrenologist Johann Spurzheim (1776-1832). Physiologist Pierre Flourens, who is the founder of experimental brain science in the early 1800s, investigated functionalities of different regions of the brain [22].

The emergence of x-ray computed tomography (CT) in the 1970s gave clinicians the opportunity to observe various characteristics of the brain without the need for surgery. Electroencephalography (EEG), which measures electrical signals in the scalp in response to a stimulus, has opened up new possibilities for studying brain function [23]. Positron emission tomography (PET) is the first technique to allow the creation of maps of the brain by measuring blood flow while the subject is performing a cognitive task. During the same period, another technique, called Magnetic Reso-

nance Imaging (MRI), that promised better pictures of the brain was being developed. MRI generates visual information of various tissues with high contrast [24]. The pioneers P. Lauterbur and P. Mansfield of MRI technique received the Nobel prize for medicine in 2003.

Brain imaging has been undergoing a revolution in the past decade with fast development, more accurate and less invasive devices. Although invasive techniques have a better spatial resolution, neuroimaging allows fast, repeatable and multi-mode measurements of structure and function in the human brain. The capabilities of brain imaging techniques offers new insights into neuroscience, neurology, psychiatry, psychology, and even contributes to the philosophical debate about the relationship between the mind and the brain [23].

2.2.1 Neuroimaging Techniques

Neuroimaging includes various techniques for imaging the structure or function of the brain. The brain imaging techniques fall into two broad categories [25].

- **Structural imaging** deals with the anatomic structure of the brain. It is widely used for the diagnosis of large-scale intracranial diseases, such as tumors. The widely used structural imaging techniques are computed tomography (CT), X-ray, and magnetic resonance imaging (MRI).
- **Functional imaging** enables the researchers measuring the dynamic functionality in the brain regions. It allows simultaneous observation of brain's neural activity. It enables direct visualization of information processing in the brain, as the activity in the relevant area of the brain increases metabolism and is detected in the scan. The common techniques are positron emission tomography (PET), single-photon emission tomography (SPECT), functional magnetic resonance imaging (fMRI), functional near-infrared spectroscopy (fNIRS), Electroencephalography (EEG), and Magnetoencephalography (MEG). However, the most used imaging techniques in functional brain imaging are fMRI and EEG.

Functional neuroimaging modalities can be categorized according to temporal resolution and spatial resolution abilities. The temporal resolution determines the capability to separate brain events in time. It refers to the closeness of the measured activity to the timing of actual neuronal activity. The spatial resolution determines the capability to discriminate changes in an image across spatial locations. It refers to the accurate localization of the activity measured in the brain.

Figure 2.2 shows a spatial and temporal comparison of the most common functional neuroimaging modalities [2]. The modalities placed near the bottom of Figure 2.2 have better spatial resolution than those above. The techniques shown on the left of the figure have better temporal resolution than those on the right. Therefore, modali-

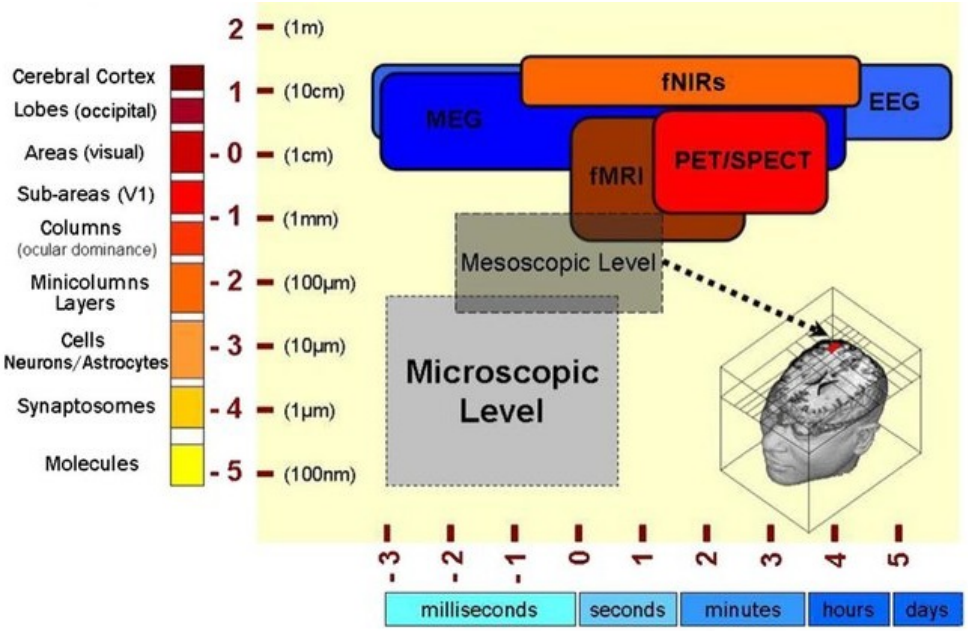


Figure 2.2: A spatial and temporal comparison of the functional neuroimaging modalities [2].

ties such as EEG and MEG have very good temporal resolution, while fMRI and PET have better spatial resolution.

Although a single imaging modality does not serve all the purposes in functional brain activity research, each has unique advantages and limitations. Deciding which technique and tool to use depends on the particular research questions to be answered. However, among these methods, fMRI stands out with its widespread use. It has been reported as a promising tool for evaluating brain functions and neurodegenerative conditions, such as parkinson's disease, huntington's disease, alzheimer's disease, and autism spectrum disorder [26].

2.2.1.1 Functional Magnetic Resonance Imaging (fMRI)

fMRI allows researchers to noninvasively measure neural activation spatially and temporally. It enables to observe the activity of the brain, including subcortical structures. This method is the most widely used brain imaging technique to investigate the living human brain while subjects perform tasks and experience mental states. Digitally enhanced MRI images of the human brain are shown in Figure 2.3.

fMRI technique detects oxygenation and flow changes in the blood that occur in response to neural activity. When a brain region is activated, it consumes more oxygen. As a result, blood flow in the active area increases to meet this demand. Through a process called "hemodynamic response", the blood releases oxygen to active neurons at a higher rate than inactive ones. This causes a change in the levels of oxyhe-

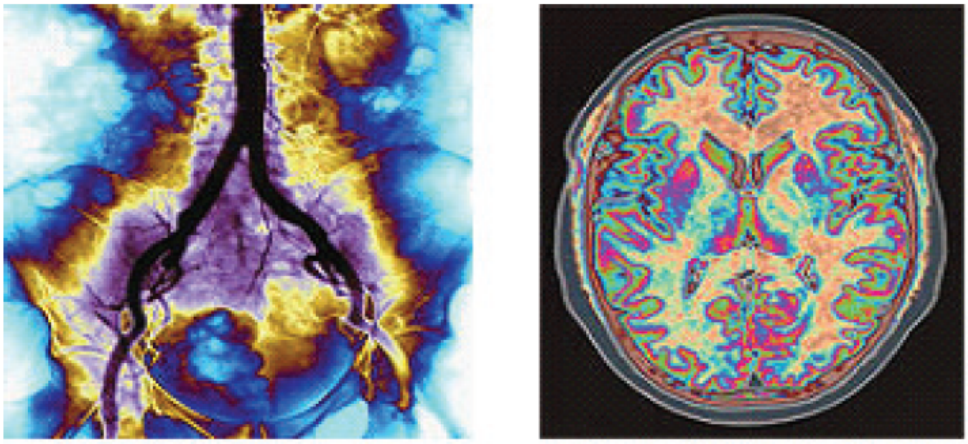


Figure 2.3: Digitally enhanced MRI images of the human brain [34].

moglobin and deoxyhemoglobin in blood, which can be detected by the degree of magnetic susceptibility [27].

In an fMRI session, the participant lies down in the scanner. It is asked to perform a specific task that is intended to be examined. Typically, the experimental task is presented on a computer screen that participants view through a mirror located in the scanner. Then, the area of activity can be detected through blood flow from one part of the brain to another by taking pictures about less than a second apart, showing where the brain "lights up" [28]. fMRI recordings enable us to observe the time instance of an activation together with the correspondence of relevant regions, which work together to generate a cognitive process.

Compared to EEG and MEG, fMRI has better spatial resolution. The main disadvantage of this technique is relatively low temporal resolution, due to the rather slow hemodynamic response to neuronal activity.

Advanced techniques and new generation fMRI technologies are developed to improve the temporal resolution. Multiple coils can be used to speed up acquisition time. Ultra-high resolution MRI spectroscopy works at tens of micrometers spatial resolutions. In recent years, ultra-high field (UHF) MRI has enabled imaging in an MRI scanner with a main magnetic field strength of 7 teslas or greater. Higher magnetic fields dramatically increase MRI sensitivity [29]. Also, rapid fMRI has been shown to detect rapid brain oscillations. The researchers were able to view brain activity that fluctuates rapidly during a cognitive process by significantly increasing the fMRI scanning rate [30].

Recently, the simultaneous acquisition of fMRI and EEG is preferred as a potentially powerful multimodal imaging technique for measuring the functional activities of brain [25].

2.2.1.2 Electroencephalography (EEG) and Magnetoencephalography (MEG)

Electroencephalography (EEG) technique records electrical activity in the brain by measuring voltage fluctuations of the ionic current within neurons [31]. This electrical activity is captured by multiple electrodes placed on the scalp. EEG was introduced long before fMRI. It has been used since the 1930s. Applications of this technique were reported as early as 1978 [32]. In contrast to fMRI, EEG is a direct reflection of neuronal function. It measures variations of electric potential on the scalp, in millisecond resolution, in line with the timescale of neuronal synaptic activity.

While EEG records the electrical activity in the brain, a newer technique, **Magnetoencephalography (MEG)**, records the magnetic field generated by this electrical activity [33]. Thus, MEG and EEG measure two complementary activities in the same neural sources. Since magnetic signals emitted by firing neurons do not need to be conducted to the scalp, measuring them by MEG enables relatively greater localization than measuring the electric currents [23, 24]. Localization is relating mental events to the locations of neural signals to find areas of the brain that are active during certain tasks or behaviors. Like fMRI, a scanning machine is used for MEG, but unlike fMRI, it does not emit radiation or magnetic fields. MEG enables us to localize the signals with a better temporal resolution of neuronal events, compared to EEG. However, MEG is costly and its ability to accurately detect events in subcortical structures is limited [23].

Although both EEG and MEG provides us powerful and insightful neuroimaging techniques with a high temporal resolution, they are not well suited for studies in which precise functional localization is important. The major drawback of EEG is that we can only record the activities related to the cortex. It is not possible to directly measure the neural activity in subcortical structures. Localizing active brain areas from EEG requires solving a difficult inverse problem which is the challenge of identifying the position of the current sources from electrode potentials [34]. The source of the signal can be very far from the point, where the scalp is measured and is affected by factors, such as head shape and dipole position and orientation; therefore it is necessary to employ source localization algorithms to determine the source of a signal [33]. The resulting spatial resolution within the brain is of the order of a centimeter for EEG and MEG [34]. However since the EEG technology is inexpensive and safe, it is widely used in studies, where the other scanning techniques are not applicaple, such as continuous monitoring during sleep [23].

2.2.1.3 Functional Near-Infrared Spectroscopy (fNIRS)

Functional near-infrared spectroscopy (fNIRS) is a portable, non-invasive functional neuroimaging technique. It captures the changes in optical properties of brain tissue [35]. This method is based on the changes in absorption of light emitted by sources onto the surface of the head. It measures hemodynamic responses by near-infrared light propagating through the head by recording the signals about the volume,

oxygenation and blood flow. A sensor is attached to the subject's forehead and connects to a computer or portable computing device, which records signals while the subject performing the given tasks [36]. fNIRS can acquire data rapidly and generate three-dimensional spatial image without providing the anatomical structures. fNIRS technique allows the design of portable, noninvasive, and minimally intrusive monitoring systems with a relatively high temporal resolution (milliseconds). However, it has a rather low spatial resolution (centimeters) compared to fMRI, EEG and MEG. Furthermore, it cannot be used to measure cortical activities more than 4 cm deep.

2.2.1.4 Positron Emission Tomography (PET)

Positron emission tomography (PET) measures glucose levels in the brain to capture neural firings. It produces a three-dimensional image of functional processes in the brain. However, PET is a nuclear medicine imaging technique that requires the patient to receive a small injection of radioactive material. The image is generated measuring the penetration of this radioactive material in the active areas of the brain. PET scans locate the areas of brain activity in the millimeter range, with a low temporal resolution of 5-10 seconds. Also, PET scans are costly. However, they can be used for medical diagnosis as they can be useful in monitoring visual problems, tumors and metabolic processes.

Single-photon emission computed tomography (SPECT) is a nuclear medicine tomographic imaging technique. This technique records the signals from gamma rays, using synchronized gamma cameras. The multiple 2-D images are reconstructed to 3-D images [37]. A section of brain can be examined from multiple angles, but it is a little less clear than a PET image. The SPECT scanners are less expensive than the PET scanner. They use radioisotopes that last longer and are easier to obtain. Monitoring blood flow in the brain determines the location of the metabolic activity. However, SPECT has problems such as long scan times and low-resolution images prone to artifacts and attenuation. Unlike the PET scans, it does not provide a quantifiable estimate of the blood flow.

Figure 2.4. listed the main advantages and disadvantages of neurocognitive brain imaging techniques, commonly used in neuroscience research [3].

2.3 Decoding Functional Brain Imaging Data

Nowadays, the amount of data collected by neuroimaging methods is exponentially increasing. Therefore, analyzing the big datasets of multidimensional fMRI recordings in a mathematical framework is critical to ensure the correct integration and comparison of the collected information.

One way of dealing with big data problems is to partition the brain into different functional regions. Then, brain decoding methods investigate the active regions to estimate the mappings between the subtle patterns to a particular event, task or thoughts.

Tech.	Advantages	Disadvantages
fMRI	• Very high spatial resolution (milimeters) • Whole brain coverage • Structural and functional data • Good source localization	• Relatively low temporal resolution (sec) • Sensitive to motion artifacts • Constraints on body position • Expensive
fNIRS	• Relatively high spatial resolution (milliseconds) • Recording in natural body positions • Low sensitivity to motion artifacts • Portable • inexpensive	• Low spatial resolution (centimeters) • Only cortical brain coverage • Influence of extra-cerebral hemodynamics • Influence of hair and skull characteristics
PET	• High spatial resolution (milimeters) • Whole brain coverage • Metabolic data	• low temporal resolution (sec) • injection of radioactive tracer • expensive
EEG	• Very high spatial resolution (milimeters) • Portable • inexpensive	• Low spatial resolution • Sensitivity to environmental noise • Inverse problem of source localization • Time-consuming preparation
MEG	• High temporal resolution (milliseconds) • High spatial resolution (milimeters) • Good source localization	• Sensitivity to environmental noise • Contraindications (e.g. dental crowns) • Non-portable • expensive

Figure 2.4: The main advantages and disadvantages of neurocognitive brain imaging techniques that are most commonly used in neuroscience research [3].

It is possible to create functional maps of the brain using the data collected from neuroimaging tools. The mapping process intends to establish a systematic relationship between the patterns of brain activity and the external or internal stimuli. In order to study this relationships, the data gathered from many subjects are aggregated under a mathematical model, which link brain anatomy to the distribution of cognitive function across the brain.

2.3.1 fMRI Data Analysis Process

A standard fMRI database contains BOLD signal time series, recorded at multiple voxels in the brain [38]. A **voxel** is a three-dimensional cube created by MRI scanning software to partition the brain volume into a grid. A high-resolution fMRI brain scan produces voxels about 1 cubic millimeter in size, which summarizes the activity of about a hundred thousand neurons in about 1 second. The key aspect of fMRI is that information is not distributed uniformly across voxels, but rather some groups of voxels (e.g., those corresponding to a specific anatomical region) are more informative for a particular task than the other groups [39].

When conducting a study of brain function using fMRI, the image of the brain is scanned repeatedly while the subject is presented with a stimulus or asked to perform some task. The fMRI data contains consecutive brain volumes, where each volume

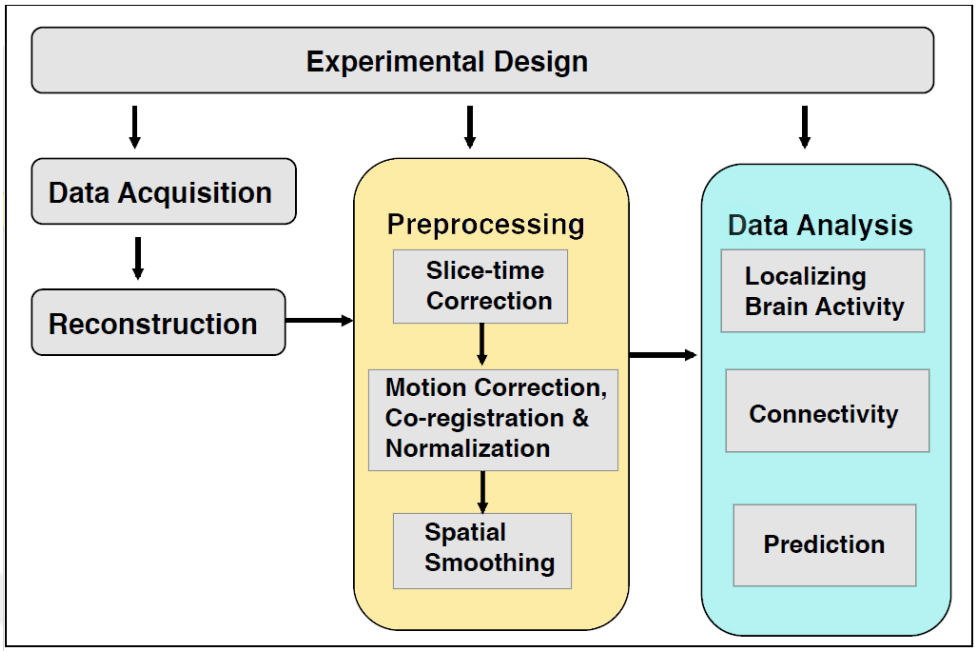


Figure 2.5: The fMRI data processing pipeline that illustrates the different steps involved in a standard fMRI experiment [62].

is a three-dimensional collection of voxels that capture a one-time frame of brain activity. However, the data directly from the scanner is not very suitable for identifying brain areas activated by the task under investigation.

Figure 2.5 shows the basic steps of the fMRI data processing pipeline accompanying a standard fMRI experiment. As it can be seen from this figure, the initial step is the design of the fMRI experiment, followed by the acquisition of brain activation data. Then, the fMRI data goes through a series of preprocessing stages necessary to organize the data for targeted task-related analysis. Preprocessing aims to minimize the effect of data acquisition and physiological artifacts, and standardize the locations of brain regions among subjects to improve experimental analysis. The major steps involved in fMRI preprocessing are slice timing correction, motion correction, coregistration of structural and functional images, normalization and smoothing. The details of each preprocessing step are explained in [14].

After data acquisition and data preprocessing is completed, fMRI data of a single participant is organized in a spatiotemporal format of voxel-specific MR signal time series. An example of these spatially arranged voxel time series are shown in Figure 2.6., where rows represent time and columns represent spatial coordinate index of voxels.

fMRI experiments are often repeated for several runs in the same session; several sessions on the same subject and several subjects drawn from a population. Multi-session/subject experiments allow us to determine whether the observed effects are common and stable across or between groups for generalization of the conclusions.

	Voxel 1	Voxel 2	Voxel 3	Voxel 4	.....	Voxel n_V
Time-point (image) 1	134.81	136.70	133.87	128.69	.....	142.83
Time-point (image) 2	136.70	138.58	135.28	130.10	.....	145.18
Time-point (image) 3	135.76	137.17	134.81	129.63	.....	143.30
⋮	⋮	⋮	⋮	⋮	.....	⋮
Time-point (image) n_T	133.20	135.56	132.23	131.18	.....	142.11

Figure 2.6: An array representation of voxel-specific MR signal time-series of a single participant and a single fMRI experimental run. Rows represent data acquisition times and columns represent voxels, i.e., three-dimensional locations. n_T and n_V denote the total number of time-series and total number of voxels, respectively.

In general, a two-level hierarchical analysis is made, within-subject (individual) and across-subject (group).

2.3.2 Univariate and Multivariate Analysis Approaches

A standard fMRI study gives rise to massive amounts of noisy data with a complicated spatiotemporal correlation structure. Since the data is difficult to interpret, a variety of statistical techniques are used in the analysis of the fMRI data. These techniques include basic statistical tests, such as t-test or Kolmogorov–Smirnov test, frequency domain analysis methods, such as, Fourier or Wavelet analysis, pattern classification methods, such as, artificial neural networks or graph theory [40].

A wide range of available techniques for analyzing the neural activities in the brain, can be grouped under two major approaches:

- Univariate approach, based on analyzing the time series of each voxel independently,
- Multi-voxel pattern analysis approach, which aggregates the time series of all voxels.

Conventional techniques, such as, independent linear models are mostly bases on univariate fMRI data analysis. In this group of univariate models, each voxel is mapped to a cognitive function. In other words, these methods estimate the relationship between cognitive variables and individual brain voxels or regions, averaged over trials. A common tool of univariate voxel analysis approach is statistical inference applied on General Linear model (GLM) to detect active brain regions by searching for linear correlations between the fMRI time course and a reference model defined by the experimenter [41].

In 2001, Haxby et al. show that information could be decoded from patterns of fMRI activation across voxels [42]. This pioneering study had a strong influence on the neuroscience community. After that, a class of techniques referred to as *Multi-Voxel Pattern Analysis (MVPA)* has been applied to a variety of studies [43–45].

MVPA techniques analyze multiple voxels simultaneously using the brain data. The simultaneous analysis of multiple voxels can thus pick up on patterns across brain space. Since it focuses on the analysis of distributed patterns of neural activity, it can detect differences between conditions with higher precision than traditional univariate analysis. The increased precision afforded by MVPA methods makes it feasible to measure the presence/absence of cognitive states based on only a few seconds of brain activity. The main benefit of analyzing multiple voxels at the same time is that it takes into account spatial interdependencies across voxels. Each voxel's response is analyzed for controlling the other voxels in a data set. Common choices of voxel sets are spherical region of interest (searchlight), anatomical regions of interest, or whole brain (all gray-matter voxels).

There are many MVPA techniques developed to classify mental states or characterize the representation of the information encoded in the brain [46]. These techniques analyze recorded activity patterns using *machine learning* techniques to detect the information coded within populations of voxels.

Their algorithms "learn" a functional relationship between brain response patterns and a subject's cognitive state expressed by a label. This learned functional relationship is then used to predict unknown labels from fMRI data [47].

2.3.3 Major fMRI Data Analysis Methods

A large number of fMRI data analysis methods have been developed to extract meaningful information from the BOLD signals. In this part, we will review the most common ones. The main purpose of these methods is to determine how mental representations map to neural activity patterns, to qualitatively and quantitatively evaluate the activated regions The methods that are used to decode fMRI measurements are divided into two broad classes:

1. Model-driven methods.
2. Data-driven methods.

In the following subsections, we shall briefly overview these methods.

2.3.3.1 Model-driven Methods

Model-driven methods, which are more traditional than data-driven methods, are widely used because of the easiness of application and simple interpretation. These methods usually rely on predefined seed regions or voxels.

They select certain regions of interest (ROI) as seeds, then, correlates the ROI and spatio-temporal BOLD signals from the rest of the brain. They select specific regions (ROI) as seeds, then correlate ROI and BOLD signals from the rest of the brain. They, then, determine whether other regions are associated with the selected seeds according to predefined measures.

The advantage of these methods for analyzing the fMRI BOLD signal is that the results are relatively simple to interpret because of the focus on specific ROIs. However, the resulting inferences depend on the selection of initial seeds. Different seed selections may lead to different results. Therefore, the selection of a priori ROI requires skill in neuroscientific background. Another drawback of these methods might be that they can only investigate the patterns that are already predicted or known. Thus, they may not detect unexpected patterns which are not included in the models [38].

A commonly used model-driven method is **Cross-Correlation Analysis (CCA)**. It, suggested by Biswal et al (1995) [48], was central to the discovery of functional connectivity of MRI. Using this method, a correlation can be estimated between pairs of functionally connected anatomical regions. This method can be considered as the extended version of **Pearson correlation coefficient** [49]. Cross-correlation calculates the linear correlation between all possible shifted versions of an fMRI signal relative to the other signal.

Pearson correlation only reflects the marginal association between the voxel time series. Therefore, it may not be seen as a suitable tool to capture the direct functional link between them [38]. A large correlation can be detected between a pair of fMRI signals because of their common correlation with another signal [50]. This method often identifies significant connections. Yet, it is difficult to distinguish which of these correlations reflect actual connections and which are due to other factors [50].

Sun et al. [51] developed an alternative metric for correlating the voxel time series using **Coherence Analysis**, by applying the correlation concepts in the frequency domain. Since the coherence of the voxel time series is not affected by frequency shifts, it is insensitive to regional differences in blood flow and volume [38]. Coherence Analysis estimates a correlation between two BOLD responses, in the frequency domain. The coherence between two BOLD responses are approximately equal or proportional to the coherence between the neural activations that elicited those BOLD responses, even if the two regions are characterized by different hemodynamic response functions [52].

When considering multiple seed anatomical regions, the specific contribution of each functional link to each region needs to be determined. **Partial Correlation (PC)** is a suitable technique for such situations. It determines the functional connectivity

between two specific regions while removing the influence of all other factors [38]. However, the application of partial correlation in investigating brain connectivity, especially in large-scale brain networks, is limited due to the technical challenges in its estimation [50].

Another model-based method for functional connectivity is **Statistical Parametric Mapping (SPM)**, created by Karl Friston in 1996 [53]. It combines General Linear Model (GLM) and Gaussian Random Field (GRF) theory to provide functional connectivity between spatially extended data. [38].

Regional Homogeneity Method (ReHo), suggested by Zang et al. in 2004 , evaluates brain activity by synchronization of the time series of underlying voxel cluster with its nearest neighbors [54]. In this method, no priori ROI selection is required for assessing the intrinsic brain activity. Kendall's coefficient of concordance technique [54] is used to evaluate the time series similarity within a cluster of a specific voxel and its nearest neighbors. Compared to other model-driven methods, ReHo analysis seems less sensitive to hemodynamic response under active conditions. However, this method can detect unpredictable patterns in hemodynamic response, as it does not require a prior ROI selection. [26].

2.3.3.2 Data-Driven methods

Unlike regression-based hypothesis-oriented analysis methods, the main advantage of data-driven methods is that they can be applied to experimental paradigms without a priori model of brain activity. They are also called exploratory data analysis methods. Since they are not based on a priory model, they provide the opportunity to identify unpredictable correlations in the data [38]

Most common data-driven computational methods can be classified as:

1. Decomposition-based methods.
2. Clustering analysis methods.

Let us briefly overview the above data-driven methods.

2.3.3.2.1 Decomposition-Based Methods

In this group of methods, the goal is to express the original fMRI dataset as a linear combination of basis vectors or that of statistically independent components [55]. The main decomposition-based methods are Principal Component Analysis (PCA) [56] and Independent Component Analysis (ICA) [57].

Principal Component Analysis (PCA) is a popular method from multivariate statistics. The idea behind this method is to display the signal using a linear combination of several orthogonal contributors. PCA decomposes each contributor into a temporal

pattern, which is called the principal component, and a spatial pattern which is called an Eigen map [58]. Basically, PCA orthogonally rotates data to a new coordinate system. It finds the direction in the data along the maximal variance, then, rotates so that this direction becomes a coordinate axis. PCA is assured to be optimal only if the underlying data have a multivariate normal distribution. However, this may not be a valid assumption for the fMRI data [59].

Independent Component Analysis (ICA) is one of the popular methods for analyzing multivariate fMRI data. It decomposes a signal into subcomponents that are statistically independent and non-Gaussian [60]. Both ICA and PCA are multivariate techniques which decompose the BOLD time-series into a set of independent components [59]. They assume that the components are statistically independent. However ICA does not assure the orthogonality of independent components. Also, ICA assumes that components are non-normally distributed. This is a critical assumption; because, ICA is guaranteed to fail on data, constructed from normally distributed components.

ICA does not require a priori definition of seed regions (spatially and temporally), but unlike PCA, it analyses the entire BOLD time-series and decomposes them into independent components. However, the independent components have to be selected manually [26].

2.3.3.2.2 Clustering Analysis Methods

Clustering methods are used to group the fMRI BOLD responses into "similar" subsets. The measure of similarity is selected depending on the neuroscientific goal. In most of the problems, Euclidean distance is measured between the voxels or voxel time series. The most popular clustering analysis methods employed in data-driven methods include K-means clustering, fuzzy K-means, hierarchical clustering, partitional clustering, support vector machines, spectral clustering and graph based clustering .

K-means Clustering method partitions the sample dataset into k groups, where each data point is placed in the group with the closest mean. It begins with initial estimates of cluster means; iteratively refines the cluster means by assigning each datum to its closest cluster. Then, the algorithm updates the cluster means based on the new assignments.

Fuzzy K-means is a clustering method that allows fuzzy partition of the dataset [38]. It aims to minimize an objective function typically defined as the total distance between all patterns and their cluster centers [55]. These distances describe the correlation degree between the two fMRI signals. Brain regions whose distance is under a certain threshold are taken as functionally connected [55]. In both K-means and fuzzy K-means algorithm, selection of number of clusters is a crucial preliminary step, which has a significant effect on the neuroscientific findings [61].

Hierarchical Clustering is used to improve the model-based analysis method utilizing a correlation matrix built from the multiple seeds to resolve which regions are most closely connected [26]. Different from FCA which uses an empirically chosen number of initial clusters, hierarchical clustering analysis initially considers each voxel as one cluster, then merges the nearby clusters based on distance measurement [55].

Support Vector Machine (SVM) [62–64] is a supervised machine learning model, widely used supervised classifiers of fMRI data due to their high performance, their ability to deal with large high-dimensional datasets, and their flexibility in modeling diverse sources of data [47, 65, 66].

The clustering analysis methods can also be thought of as feature generation techniques.

A fundamental prerequisite for these techniques is a robust and reliable feature extraction method. Then, they are used for classification. In general, the success of a classifier is based on the quality and relevance of the features provided to fulfill the discrimination process [38].

2.3.3.2.3 Graph Based Analysis

Graph theory offers a theoretical framework to examine brain connectivity. Recently, there is a rapid growth in studies applying graph theory to reveal characteristics of the organization of functional brain networks.

In most of the graph based methods, the region of interest (ROI) corresponds to **nodes** and the correlation between the ROIs corresponds to the connectivity of the **edges**. Estimating the level of functional connectivity between all possible pairs of nodes or detecting the presence of a functional link using a statistical threshold provides a graphical representation of the functional brain network. Figure 2.7 depicts the fMRI data analysis steps used for extracting the brain graphs. This graphical representation enables the analysis of brain network organization by graph theory. It describes different brain networks using the key properties of graph theory such as clustering coefficient, node degree, betweenness and closeness centrality, "small worldness" and "rich club" indices. These network features depict organizational properties in the relationships of voxels or regions with others.

There are various applications of fMRI data analysis using graphical representations. Zhang et al. (2011) [67] applied the social network theory to resting-state fMRI data to identify regions within the epileptogenic network. They observe that the model could reveal abnormal network properties. They separated the medial temporal lobe epilepsy patients from normal control subjects by a classification algorithm using network data. Khazaee et al. (2015) [68] used graph theoretical methods with machine learning to study functional brain network changes in Alzheimer's disease patients. They used a SVM classifier for assessing graph measures in the diagnosis of the dis-

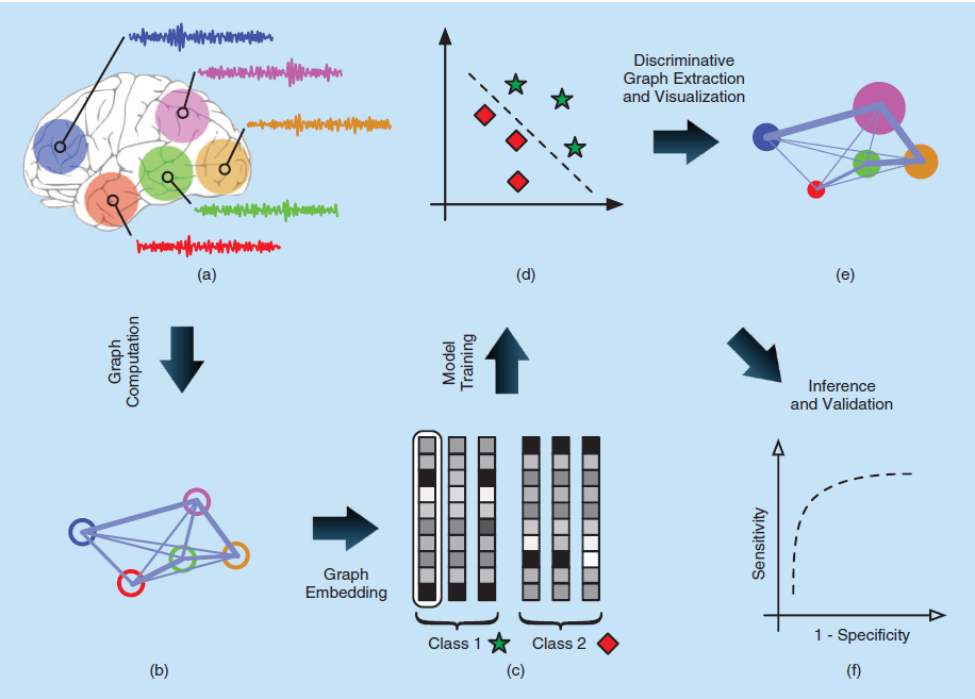


Figure 2.7: fMRI data analysis with brain graphs. (a) First, the imaging data is preprocessed. Then, the brain is partitioned into regions, mostly using a brain atlas. Each region is assigned a regional representative time series. (b) A labeled simple graph is generated from the regional time series, where edge labels correspond to statistical dependency between brain regions, and brain regions are mapped to graph vertices. (c) The graph is embedded into a vector space. (d) statistical machine learning can be used for decoding the cognitive states. (e) Graph statistics can be estimated to analyze the brain graph [4]].

ease. They show that the graph theory can describe different aspects of the brain network, while also providing promising results for examining changes in connectivity among brain regions in disease patients.

A number of studies proposed **Mesh Network Models** to estimate the relationships among nodes within a local neighborhood. Ozay et al. (2012) [69] propose **Local Mesh Model** (LMM) to model the relationship among the spatially neighboring voxels. They show that features of LMM perform better than the raw voxel intensity values for the classification of cognitive states. Firat et al. (2013) [70] propose a **Functional Mesh Model** (FMM), in which local meshes are formed around nodes by selecting the neighbors within a functional neighborhood. Onal Ertugrul et al. (2018) proposed **Hierarchical Multi-resolution Mesh Networks** (HMMNs), which forms a set of brain networks at different time resolutions of BOLD time-series to represent the underlying mind states. The fMRI signal is decomposed into a few frequency subbands; then, a brain network is generated with a set of local meshes at each subband. "The zone around each anatomical region is defined by a neighborhood system based on functional connectivity. The arc weights of a mesh are defined by ridge regression formed among the average region time series" [71]. Then, the adjacency matrices for the networks generated at different subbands are constituted.

The use of graph theory requires careful selection of nodes. The estimated networks will be only as accurate as the nodes. Defining nodes is especially difficult, because nodes may not the same across the samples. Another problem with graph representations is that number of edges tends to grow exponentially if a typical machine learning algorithm of moderate complexity is applied. For this reason, many state-of-the-art graph algorithms can deal only with graphs including some hundred up to a few thousand nodes at maximum. The estimation of dependencies is rather difficult since the number of vertices is large compared to the number of points in space/time.

2.3.3.2.4 Artificial Neural Networks

Artificial neural networks (ANN) have recently gained importance in fMRI analysis for supervised prediction and classification using **deep learning** methods. Inspired by the architecture of the brain, artificial neural networks include layers of feature extraction units called artificial neurons. These units learn multiple levels of abstraction directly from the data [72].

In the last decade, many available artificial neural network models are applied for the analysis of fMRI data. For example, **Radial Basis Function Networks** (RBFN) were used to automatic diagnosis of Autism Spectrum Disorder [73], **Feedforward Neural Networks** (FNN) were used to classify Schizophrenia [74] and **Convolutional Neural Networks** (CNN) were used to diagnose amnestic Mild Cognitive Impairment [75].

Convolutional Neural Networks are widely used for 2-dimensional data for building an auto-encoder. Auto Encoders are specialized neural networks for generating

reduced feature sets through nonlinear input transformations. They have been used for feature reduction of functional connectivity in several studies [76–79]. Nie et al. (2016) [80] used CNN to extract features from the fMRI signal to estimate the life expectancy of patients with brain tumors. Firat et al. (2015) [81] constructed a sparse encoder in conjunction with CNN for classification to conceive the changes in brain state following memory words . Kivilcim et al. (2018) proposed an Artificial Brain Network model to capture the functional connectivity among anatomical regions during a task. "They estimate a set of brain networks, each of which represents the connectivity patterns of a cognitive process. Then, they use the edge weights of the estimated brain networks to train a Support Vector Machines classifier to label the underlying cognitive process" [82].

While artificial neural networks opens a new era in big data analysis, there are some limitations that arise for both deep learning and classical machine learning methods in using fMRI data. For example, the sample size of fMRI data is relatively smaller compared to the number of functional connectivity parameters due to a lack of labeled data. Even if the number of samples required for the training of a standard machine learning method is extremely lower than required for deep learning techniques, the small sample size can lead to overfitting problems in both cases [83].

In many studies, deep learning based classifiers are used in combination with some other feature selection or feature compression methods. This suggests that the achievement of a reliable classification accuracy requires statistically significant amount of fMRI data to train the deep models.

Successful deep learning applications for decoding the cognitive tasks from the fMRI data require a network architecture suitable for a specific task [84]. As the qualitative and quantitative properties increase in fMRI data, the application of deep learning methods for brain decoding performs better.

2.3.3.2.5 Information Theoretic Methods

Application of the information theoretic approach to functional neuroimaging data is becoming increasingly common. Shannon information theory identifies a principled and unified framework for the statistical analysis of neuroimaging data. **Shannon entropy** [5] measures the amount of information that changes in neural responses according to different stimuli.

There a number of novel neuroimaging data analysis approaches using Shannon information theory. Many studies have tested and evaluated various types of entropy measures on fMRI data. Besides the classical Shannon entropy, its different variations have been adapted by the studies included *differential entropy* (DE) [85], *permutation entropy* (PE) [86], *multiscale permutation entropy* (MPE) [87], and *sample entropy* (SampEn) [88–92].

The other variations of entropy, namely *transfer entropy* [93, 94] and *maximum entropy* [95–97] are also used to analyze fMRI data. Transfer entropy measures the directed information transfer, like Granger causality [98]. Maximum entropy is a framework that characterizes the statistics of neural activities using the maximum entropy principle [99]. Although there are a variety of entropy measurements, there is a basic computational similarity among these measures and the original Shannon entropy [12].

Mutual Information (MI) [5] is another information-theoretic measure, which is used for brain connectivity analysis [100–102]. It can be effectively used in the study of complex networks due to the ability to flexibly capture the relationship of various forms. Mutual information measures the interdependence between two or more time series based on the similarities of their probability density functions. It is a measure of the decrease in uncertainty about a random variable that informs about another. Higher mutual information indicates a higher probability for the existence of a connection, while lower mutual information indicates that the existence of a connection is unlikely [100].

2.4 Chapter Summary

In this chapter, we provide an overview of the processing of neural information in the human brain, measuring brain activity and obtaining information about brain activity from these measurements using neuroimaging data analysis methods.

In order to understand the mechanism of processing information in the brain, we first briefly explained the structure of brain cells, the way that brain cells generate and convey neural signals, and the way these cells communicate with each other within neural circuits.

In the next section of this chapter, we described measuring neural activity with various neuroimaging techniques. We noted that these techniques fall into two broad categories as "structural imaging" and "functional imaging". While structural imaging techniques deal with the structure of the brain, functional imaging techniques are used to measure the simultaneous observation of dynamic functionality in the brain.

We, then, reviewed the most common functional imaging techniques which are PET, SPECT, fMRI, fNIRS, EEG, and MEG. We made a comparison of these functional neuroimaging modalities with respect to spatial and functional characteristics. We observed that that frequently used imaging techniques in functional brain imaging are fMRI and EEG. However, the decision on which technique and tool is appropriate depends on the research goals and methods.

Finally, we explained the methods used to decode and analyze brain activity patterns in brain imaging data. In many research studies, the brain is partitioned into different functional areas. Data collected from functional neuroimaging tools are investigated to find the systematic relationship between the patterns of brain activity in these areas

and the external or internal stimuli. In this context, we first explained the data processing steps involved in a standard fMRI experiment. Then, we briefly reviewed two common approaches for analyzing fMRI data that are voxel-wise univariate approach and multi-voxel pattern analysis approach. Finally, we described the most common model-driven and data-driven fMRI data analysis methods.

CHAPTER 3

AN INFORMATION THEORETIC APPROACH FOR MODELING COGNITIVE STATES USING FMRI IMAGES

> "The ultimate purpose of life, mind, and human striving: to deploy energy and information to fight back the tide of entropy and carve out refuges of beneficial order."
>
> Steven Pinker

Information theory has been partially used for neuroscientific research for many years [88,89,94,103–107]. Information-theoretic methods provide measures that are useful to understand how neuronal interactions shape the way in which neural populations represent and transmit messages about the external and internal environment.

In this chapter, we describe the basic concept of information theory which can be applied to the investigation of the dynamics of brain activity associated with cognitive processes. We also discuss the rationale behind using an information-theoretic approach for the analysis of fMRI data. Finally, we provide examples of various information-theoretic applications which have been carried out for analyzing neuroimaging data.

We believe that the suggested approach may open a door to investigate the relationship between the activities of human brain and mind.

3.1 Measuring the Information by Shannon Information Theory

"Shannon Information Theory is a mathematical representation of the conditions and parameters affecting the transmission and processing of information" [108]. It is about how to measure, represent, and communicate information effectively. The first studies on this subject were made by Harry Nyquist and Ralph Hartley, in the 1920s. However, it was *Claude Shannon*, who founded the discipline of information theory in 1948 with his famous article "A Mathematical Theory of Communication". Shannon's pioneering work presented many important ideas that guided engineers and

scientists from the 1950s. Since then the principles of classical information theory were applied to many fields.

The theory is a general framework for quantifying the ability of a coding scheme or a communication channel to convey *information*. It is assumed that the code involves several symbols (such as letters, words or neuronal responses). It is also assumed that coding and transmission processes are stochastic and noisy [109]. Note that, Shannon's information definition is associated with a symbol alphabet for encoding a message. It is not intended to relate to the meaning of the messages that they transmit.

3.1.1 Shannon Model of Communication

Shannon's communication model was mainly related to the "technical problem" of information transmission in terms of the statistical rarity of a signal. The basic information-theoretic model of a communication system proposed by Shannon is indicated schematically in Figure 3.1 [5].

The model consists of five parts:

- An **information source** produces a message or sequence of messages.
- A **transmitter** changes this message into the signal that is suitable for transmission through the channel.
- The **channel** is the medium that allows transmitting the signal from the transmitter to the receiver. During transmission the signal might be perturbed by noise.
- The **receiver** decodes the signal to recover the transmitted message from the signal.
- The **destination** is the thing for whom the message is intended.

Shannon's innovation was to model the information transmitted by a signal through a channel in terms of the total set of alternative possible messages that could be transmitted. This approach allowed a new mathematical representation of the amount of information transmitted by a particular message, the capacity of an information channel, and the effects of noise on the channel. This mathematical model of communication was identified as a promising model of human communication.

The model of communication mentioned above is well suited for modeling brain activities corresponding to cognitive processes. Loosely speaking, we can define the human brain as an information processing and communication system establishing an analogy with Shannon's communication system, as follows:

- The information source can be considered as massively parallel neurons, which are activated by external or internal stimuli, producing signals or a sequence of signals.

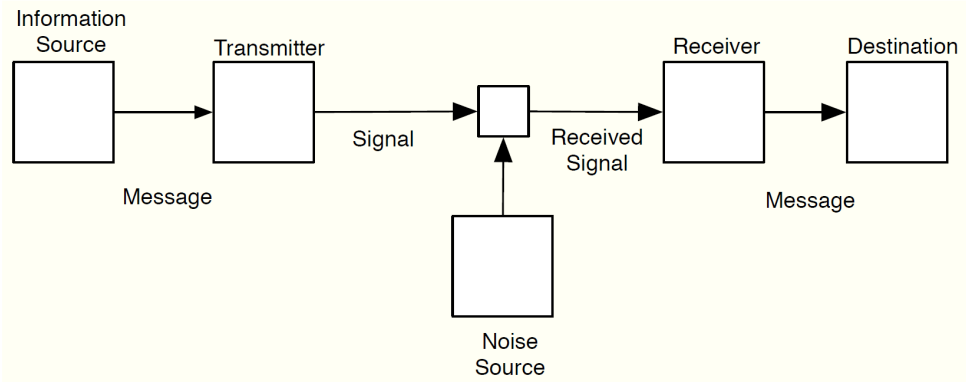


Figure 3.1: The information-theoretic model of a communication system [5].

- The activation signals generated by billions of neurons are encoded in a series of complicated electrochemical processes to protect the information contained within them.
- The encoded signals are carried along the neuron's axon. The arrival of the encoded signals causes a transmitter to be released at the axon terminal. The synapses allow the electrical signal to pass from the source to the receivers, which are the dendrites of the target neurons.
- The received signals are decoded and transmitted to the cell body as a message.
- If the total input from the dendrites to the cell body exceeds the neuron's threshold, the stimulus is viewed as the estimated information, and the neuron decides how to react.

3.1.2 Entropy in Thermodynamics and Information Theory

Shannon entropy definition is related to the idea of entropy from physics by analogy. Both concepts are concerned with disorder and uncertainty which are closely related.

The first scientific concept of entropy was introduced in thermodynamics, as a measure of disorder. It was defined by the thermodynamics' second law, introduced by Ludwig Boltzmann (1877) [110]. He developed the concept of statistical disorders by establishing statistical mechanics and thermodynamics. The second law of thermodynamics establishes the concept of entropy as a physical feature of thermodynamic systems, and interprets it as a measure of the *statistical disorder* of a system.

The close similarity of mathematical expressions of entropy in statistical thermodynamics and entropy in the information theory sometimes causes confusions. The thermodynamic entropy refers to thermodynamic probabilities, which are closely associated with the energy and temperature. On the other hand, the entropy used in the communications field is a mathematical abstraction of messages, which can be estimated from the probability distribution of the underlying random variable.

In order to avoid confusion with the concept of thermodynamic entropy, Shannon's definition is often referred to as **"Shannon entropy"** or **"information entropy"**. In this document, we will use the term "entropy" to refer to "Shannon entropy".

3.1.3 Entropy as an Average Information

The fundamental concept of information theory is **entropy**. The word is derived from the ancient Greek to mean "transformation content". Shannon introduced the mathematical measure of the average amount of information that is exchanged in the process of communication (regardless of whether the signals involved are bit streams, analog signals, or spike trains).

Shannon entropy quantifies the average amount of information required to represent an event or to specify the state of a random variable. It is important to emphasize that entropy is not associated with the value of the random variable, but depends only on the probability distribution of its values.

In a broader context, Shannon entropy can be used for measuring uncertainty about the state of a system. It represents the degree of incompleteness of our knowledge about the precise state of the system under consideration [111], namely, it is the "missing information". Therefore, a system that is likely to be in many states has high entropy.

Entropy reaches its maximum value, when the probabilities of all states are equally likely [111]. This refers to *maximum uncertainty*, *maximum information*, and *minimum knowledge*. In this case, the outcome of the system cannot be predicted. A large number of bits are necessary to describe all the possible outcomes of the system. On the other hand, zero entropy indicates that one of the states is always the outcome, but the other ones have a zero probability of being occurred.

3.1.4 Estimation of Information Content

In Shannon information theory, the entropy or the amount of information (of a value of a random variable observed) is the average or expected degree of "surprise" or "uncertainty" of the possible events [112]. The informational value of a message depends on how surprising the content of the message is. In other words, quantification of information is related to measuring how much *surprise* there is in an event or a particular outcome.

The amount of information conveyed in an event, or the level of surprise, depends on the probability of the event. If an event is very likely, it is no surprise that it occurred as expected; therefore, the message informing such an event does not contain much information. However, the lower the probability of an event occurring, the more informative the message that the event has occurred.

3.1.4.1 Self-Information (Shannon Information)

Shannon proposed a way to measure the information associated with the outcome of a random variable. Consider a probabilistic event, represented by a discrete random variable x, with possible values $x_1, x_2, \ldots, x_n$, and the associated probability mass function, $p(x)$. Self-information for the outcome $x = x_i$ is defined as:

$$I(x = x_i) = \log\left(\frac{1}{p(x = x_i)}\right) \tag{3.1}$$

or equivalently,

$$I(x = x_i) = -\log P(x = x_i), \tag{3.2}$$

where $\log$ is the logarithm, and $p(x = x_i)$ is the probability of outcome x_i. Thus, $-\log P(x = x_i)$ is considered as a measure of the degree of surprise or unexpectedness of the outcome $x = x_i$ [112].

The unit of self-information depends on the base of the logarithm. The base-2 logarithm means that the unit of the information measure is in bits (binary digits). This is the common unit of information. This can be interpreted in the information processing sense as the number of bits required to represent an event. One bit represents 2 possible states, while n bits represent 2^n possible states. If the base-10 is used, the unit is called decimal digits, or dits. In analytical work, where integration and differentiation are involved, the base-e is sometimes useful. In this case, the resulting units of information is called natural units or nats [5]. In this document, unless we want to emphasize the units, we adopt the convention prevalent in information theory of using logarithms to the base of 2 [113], and will write $log(p)$ instead of $log_2(p)$.

In the self-information formula, it is taken the *negative* of the summation since the logarithm of a number less than 1 is negative. Because probabilities are always less than one, the summation will always be negative. Thus, multiplying by - 1 changes the sign of the result to positive.

The quantity $I(x = x_i)$ is a monotonic function of the probability $p(x = x_i)$ and that expresses the information content of the outcome. It has several properties such as:

- Information is a positive quantity.
- Deterministic outcomes contain no information.
- Information content increases with decreasing probability.
- If two independent events occur, the information content will be the sum of the information gained from each of them separately.

3.1.4.2 Entropy Estimation as an Average Information

Shannon entropy is defined as the expected value of self-information. It measures the average amount of information over all possible events or outcomes within a system.

The average amount of information for a discrete random variable x, is estimated by taking the expectation of Equation 3.2 and is given by:

$$H(x) = \sum_{i=1}^{n} I(x = x_i)\, p(x) \tag{3.3}$$

or,

$$H(x) = -\sum_{i=1}^{n} p(x)\, log\, p(x) \tag{3.4}$$

Shannon called H the "entropy" of the random variable x, noting that it plays a central role in information theory as measures of information, choice, and uncertainty [5].

H will be zero if and only if all the $p(x_i)$ but one are zero; where one probability is unity (certainty). In this case, we are sure of the outcome (no uncertainty - no information). Otherwise, H is positive. H has a maximum value when all the $p(x_i)$ are equal. This is the most uncertain condition.

In the case of transmitted messages, the entropy of a message is a measure of the average amount of information transmitted by the message. For example, if a random variable x can have 8 (or 2^3) equally likely states. In order to communicate the value of x to a receiver, it is required to transmit a message of length of 3 bits, since the entropy (estimated by log_2) of this variable is:

$$H = -8 * \frac{1}{8}\, log\, \frac{1}{8} = 3\, bits/symbol. \tag{3.5}$$

According to the noiseless coding theorem [5], this is the lower bound on the number of bits required to specify the state of the random variable. When each state is equally likely, the entropy is also the number of binary questions needed to decide the content of the message [114].

The entropy in the case of two possibilities with probabilities $p(x_i)$ and (1 – $p(x_i)$), namely,

$$H = -(p(x_i)\, log\, p(x_i) + (1 - p(x_i))\, log(1 - p(x_i))) \tag{3.6}$$

is plotted in Figure 3.2 as a function of $p(x_i)$ [5].

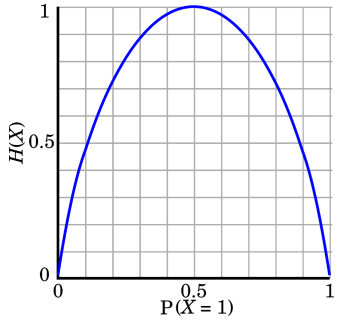


Figure 3.2: Entropy plot for two possibilities with probabilities $p(x_i)$ and $(1 - p(x_i))$.

3.2 Information-Theoretic Interpretation of Neural Activity

There is significant theoretical and empirical support for the importance of information-theoretic methods in the research of cognitive and neural processing [90, 91, 94, 115, 116]. Consequently, entropy is regarded as an appropriate measure for characterizing stochastic neurophysiological data, which is in the form of sampled real-value signals. Unlike moment statistics such as a mean and variance, it does not depend on the absolute values of the signal. Instead, it reflects the regularity of the distribution of values.

Neurophysiological systems typically exhibit complex behaviors with nonlinear dynamic properties. The application of the concept of complexity is focused on measuring regularity using various metrics for nonlinear time series analysis. These metrics are provided by the Shannon information theory [86].

The neural complexity introduced by Tononi et al., characterizes the brain in terms of local segregation and global integration during cognitive functions, where segregation means statistical independence of activities of different neuronal groups, and functional integration occurs when the neuronal groups show a high degree of statistical dependence [117]. This leads to the formulation of a measure, called neural complexity, that reflects the interplay between functional segregation and integration among the components of a neural system, measured using the concepts of Shannon entropy and mutual information [117].

The brain can also be supposed as a self-organizing system. For example, neuronal ensembles self-organize into complex spike patterns. Self-organization occurs when a spontaneous increase in order is arisen by the interactions of the elements of a complex system [118]. Generally, interactions among the components of a system (neurons or anatomical regions) induce statistical regularities and structuring information processing within the system. Shannon information theory provides a method for the measurement and interpretation of these regularities.

It is supposed that the human brain has a wider potential neurocognitive states than other species; this is an important feature of its complex behavior [119]. The nu-

merous high-probability states in the brain can give rise to distinct neuronal firing trajectories, each of which may represent a unique event or situation. Each situation activates the appropriate highest-probability state, and this matching process reflects the "best guess" of the brain [120].

Suppose a group of voxels in an anatomical region are in a certain state. When those voxels are activated by a triggered message, the message causes the voxels to go to a different state. This event can give rise to change in the state of underlying anatomical region. Measuring the entropy of this region allows us to make more accurate inferences about its state during a cognitive process.

Recently, a new theory is rapidly gaining importance in mind and brain studies. This approach describes the brain as essentially a "probabilistic prediction machine" that continually generates predictions of its actions and sensations. It is supposed to be dedicated to the task of minimizing the discrepancy between how it predicts the events to be, and the actual outcomes (Hohwy 2013; Clark 2016, 2020; Friston and Buzsaki 2016) [121–124]. This vision suggests that the brain is a prediction error minimizing device. The principal mechanism for perceptual inference is the prediction, grounded on internal, productive models, revision of model parameters, and prediction error minimization. [121]. The proponents of this theory assume that conceiving of the brain in this way has yielded the first steps towards a unified theory of mind. The probabilistic and statistical foundations of the theory are based on the Shannon information theory and statistical physics concerning entropy and mutual information.

3.2.1 Information-Theoretic Methods for fMRI Analysis

Shannon Information Theory provides a variety of tools, which can be used on fMRI data to investigate human brain activity for a given cognitive function, from neuronal groups and anatomical regions to neural networks, formed by their interconnections. It provides a probabilistic framework that quantifies the statistical non-independence between brain signals [98].

One advantage of entropy measurement over conventional methods is that it requires few assumptions about the nature of hemodynamic responses, underlying neural processes, or the data itself [104]. The information theoretic methods applied to fMRI signal features offer a model-free alternative to the study of neural information processing. The result of applying an information-theoretic measure to data is not a parameter (i.e. synaptic strength) in a model, but a value that measures some relationships within the data. It is capable of revealing both linear and nonlinear interactions. Therefore it has the potential to reveal some important fMRI signal characteristics that classical data analysis techniques ignore.

Information theory possesses metrics, such as, self-information, entropy and cross-entropy, designed to quantify the behavior of systems with any number of variables. It can be applied to any mixture of data types. When using information theoretic measurements, results are usually produced in bit units. This facilitates simple com-

parisons between voxels, anatomical regions, cognitive tasks, or subjects. The ability to measure effects in bits enables direct assessment of effect sizes [125].

Another advantage of the information-theoretic method is its ability to quantify representational interactions between different neuroimaging responses: for example, different regions, frequency bands, time periods or recording modalities [125]. Also, the neuroimaging data can first be processed by a dimensionality reduction or a wavelet transform technique, then, used for the analysis.

3.2.2 Information-Theoretic Applications in Neuroimaging Data Analysis

Recently, information theoretic approaches have become popular for the analysis of the large-scale recordings of neural activity. Studies using entropy as a representative and quantitative measure of complexity in brain dynamics enable the quantitative testing of hypotheses about encoding and transmitting the information across the brain regions [103].

Today's applications of information theory to investigate different aspects of brain dynamics range from the **information representation** [101], **neural coding** [105], **signal variability** [93, 104, 106, 107], **functional connectivity** [88, 94–96, 126], **cognitive abilities** [89, 103], **neural adaptation and behavior** [127], **categorical perception** [128], **state of consciousness** [115, 129–131], **development dynamics** of the brain [132], **aging** and **age-related brain complexity** [92, 116, 119, 133, 134] to the **cognitive impairments** [86, 91, 119, 135].

Various studies have been carried out to understand how neurons work together to represent information. For example, Samengo et al. studied information-theoretical methods that were proposed to assess the relevance of individual response features in the **neural code** [105].

The theoretical and empirical studies show that brain **signal variability** plays a crucial role in its functioning. There is ample evidence as to the underlying relations between the entropic measures and the signal variability [93, 104, 106, 107]. These findings help us to explain the crucial role of entropy in the quantification of brain information processing given the direct correspondence between the signal variance and the amount of information.

De Ajauro et al. [104] proposed a method for the analysis of event-related fMRI (ER-fMRI) based on entropy estimation using the Shannon formulation. They used entropy as a measure of the BOLD signal's variability. They divided each voxel's post-stimulus time window into two segments: the event-related signal and the baseline signal. For each time window, they discretized the possible signal values and pooled the time-dependent measurements to determine the probability of the signal. Then, the entropy for each time window was estimated. Their results show the ability of the entropy measure to distinguish between active and resting cerebral regions for motor and visual stimuli.

Lüdtke et al. [93] investigated the use of information theory to determine which changes in neurophysiological signals cause changes in the BOLD contrast, measured on fMRI. Unlike linear correlation approaches to establish covariations between neural activity and BOLD signal, they proposed transfer entropy, which can explain nonlinear causal relationships between neural activity and BOLD signal. They found that reliable and statistically significant entropy estimates can be computed even using the limited amount of data commonly afforded by empirical recordings.

Entropy manifests as a robust feature for the **classification of the mental states**. For instance, Zheng and Lu [136] proposed the use of entropy for **emotion classification** and Brasselet et al. [128] used information theory as a tool to investigate **categorical perception** mediated by representation similarity in neural activity.

Some studies applied information theoretic measures to examine various aspects of **cognitive abilities** in humans. Saxe et al. investigated the relationship between brain entropy and **human intelligence** [89]. They hypothesized that intelligence would be positively correlated with entropy. Intelligence was measured with the Shipley Vocabulary and WASI Matrix Reasoning tests using resting-state fMRI data. Brain entropy was calculated using the Brain Entropy Mapping Toolbox (BENtbx) [88], using Sample Entropy. They observed a positive correlation between individual scores on vocabulary and reasoning tasks and brain entropy in the prefrontal cortex, inferior temporal lobes, and cerebellum. They concluded that brain entropy is positively associated with intelligence.

Recently, Liu et al. used a large resting-state fMRI dataset from the Human Connectome Project to study the individual differences. They suggested that the entropy profiles of cortical activity could provide new perspectives for investigating **individual differences in behavior**. Their findings validated the whole cortical entropy profile based on fMRI as a stable marker and informative predictor of various aspects of human subjects' cognitive abilities [103].

Several studies investigated the interactions between **information uncertainty** and **cognitive control**. They explored the overlap between the neural networks associated with uncertainty and neural networks serving cognitive control [137–139].

There are many studies which employ entropy measures on quantifying the dynamics of brain activity related to **aging and cognitive functioning** as well as to the age-related loss of complexity with diseases and disorders.

McIntosh et al. investigated the relationship between **neurophysiological variability** and **behavioral variability** in maturation using Multiscale Entropy (MSE). They measured the brain signals of children and young adults during the face recognition task. Multivariate statistical analysis revealed a significant age-related increase in Multiscale Entropy. They observed that, brain signal variability increases with maturation. Subjects with higher signal variability showed less variability in response latency and higher performance accuracy. They suggested that maturation brings about differentiation and specialization of brain regions. Also, there was increased integration between distributed neuronal populations and the establishment of new

functional connections [116]. Their results showed that the observed variability in brain activity is a critical feature of its function.

Yang et al. [134] tested the hypothesis that complexity of neural activity decreases with aging and is correlated with cognitive performance in the elderly. They applied multiscale entropy (MSE) as a measure of complexity and compared the MSE of BOLD signals between younger and older groups. They examined the correlation between cognitive test scores and MSE of BOLD signals in various brain regions. They found that, older subjects had the most significant reduction in MSE of BOLD signals in the posterior cingulate gyrus and hippocampal cortex. For older subjects, the MSE of BOLD signals from default mode network areas was found to be positively correlated with basic cognitive functions such as attention, orientation, short-term memory, mental manipulation, and language. Their findings confirmed the hypothesis that the entropy of BOLD activity is correlated with aging and cognitive performance based on MSE analysis.

Jia et al. measured **complexity within the dynamic functional connectivity** by sample entropy (SampEn) to analyze its association with age. Using resting-state fMRI data, they generated a brain-wide map of sample entropy for healthy subjects. The map showed larger values in the caudate, olfactory gyrus, amygdala and hippocampus, while lower values in the primary sensorimotor and visual areas. Association analysis indicated that SampEn of the amygdala-cortical connectivity decreased with advancing age. However, this age-related loss of sampEn was not seen in schizophrenia patients. The findings suggested that SampEn of the dynamic functional connectivity is a promising indicator of normal aging [92].

Yao et al. [133] applied Shannon entropy to quantify **brain functioning at different ages** to characterize the intrinsic aging properties of the human brain. They defined it as functional entropy. Using the resting-state fMRI signals, they estimated functional entropy from correlations between different regions of the human brain. Their analysis showed that functional entropy related to brain activity increased with age. They also found significant differences between males and females. They observed that, males have lower entropy at birth than females. However, the functional entropies of the two sexes increased at different rates and intersected in about 50 years. After this age, males had higher functional entropy.

Viol et al. [115] used Shannon entropy to investigate the hypothesis that the **effect of psychedelics** is partly explained by the increased entropy of the brain's **functional connectivity**. They analyzed the fMRI data under two different conditions: in a resting state and an altered state of consciousness which occurs during the ingestion of Ayahuasca. They observed an increase in the Shannon entropy after Ayahuasca ingestion. They also found an increase in local integration and a decrease of global integration in functional brain networks.

Tononi et al. [126] introduced the **entropic functional connectivity index**, using entropy and mutual information. They observed that a particular group of brain regions can interact more strongly with each other than with the rest of the brain. They emphasized that these collaborations among specific regional groups might be important

for the determination of functional boundaries in the brain. Pritchard et al. [140] applied an entropy-based method to represent functional connectivity between brain areas. In order to demonstrate the differences between **functional brain networks** using linear and nonlinear techniques, they applied both techniques to resting-state fMRI data. They observed that in networks created using the entropy-based method, the majority of nodes have very few connections, while a few important nodes (hubs) have a huge number of connections. Entropy-based networks had a higher clustering coefficient and shorter path length than Linear Correlation-based networks.

Lizier et al. [94] proposed an information-theoretic method for brain connectivity analysis, using mutual information and transfer entropy. They intended to reveal directed functional information structure between brain regions and the structure changes according to behavioral conditions. They chose information theory as an appropriate method for multivariate analysis, which captures not only directional and nonlinear relationships, but also collective interactions. They used the fMRI time series to establish the directed information structure between brain regions involved in a visuomotor tracking task. They examined the changes in this structure as the difficulty of the tracking task is increased. They observed that the task difficulty modulates the connection strength between the cortical network regions involved in motion planning, and the cerebellum, which is involved in the fine-tuning of the motor control. They concluded that their methods would be useful in identifying interregional structure in other cognitive tasks and data modalities.

Ze Wang et al. showed that **brain entropy (BEN)** can be used to investigate brain states and brain function. Their goal was to analyze the whole brain entropy patterns for normal subjects. They first performed a series of experiments to validate the entropy measurements regarding their sensitivity, specificity, and reliability. Then, they applied sample entropy (SampEn) [90] to measure the entropy of the resting state fMRI time series and generated a 3-dimensional BEN map. They observed a sharp low-high entropy contrast between the neocortex and the rest of the brain. They found that the brain is organized into 7 hierarchical regional BEN networks that are consistent with known structural and functional brain parcellations [88].

Kant et al. [141] analyzed the EEG signals recorded during motor imagery tasks which are often used as input in **brain-computer interface** (BCI) applications. In the study, various properties of the alpha frequency band were compared by adapting the wavelet-based time-frequency analysis approach. They extracted "mean, variance, wavelet energy, Shannon entropy, log energy entropy, kurtosis and skewness" features from the EEG data recorded from the motor cortex. SVM and KNN algorithms were used to evaluate their classification accuracies. They observed that **Shannon entropy showed the highest classification accuracy** of %86.4 with the SVM classifier. The results of this study shown that Shannon entropy can be used as a promising method in brain-computer interface analysis.

Furthermore, there are many information theoretic studies on the assessment of **cognitive impairments**. B. Wang et al. [86] used resting state fMRI data to assess brain functional abnormalities in Alzheimer's Disease (AD) patients. In order to investigate the abnormal complexity of the brain, they applied **permutation entropy (PE)**

to the fMRI signals of these groups and generated their whole-brain entropy maps. The results show that AD patients have lower complexity than the mild cognitive impairment and normal controls. They concluded that complexity analysis using permutation entropy in resting state fMRI data could provide useful information about cognitive impairments in MCI and AD patients.

Considering the models of disrupted brain network connectivity in autism spectrum disorder (ASD), Maximo et al. [91] suggested that **sample entropy** would provide a novel direction for understanding brain organization. They used resting state fMRI data from 45 high-functioning children with ASD and 45 age-and-IQ-matched, typically developing (TD) children. Sample entropy was calculated for the whole brain as well as anatomical regions, using voxel values. ASD patients show significantly increased entropy in the left angular gyrus, superior parietal lobule, and right inferior temporal gyrus, while decreased entropy in the superior frontal gyrus compared to normal children. The results suggest that entropy can be used as a reliable indicator of brain dysfunction in clinical populations.

Although the above mentioned studies use different variations of entropy measure, they show that the information theoretic approaches based on Shannon entropy can be used as an effective method in brain and mind studies.

3.3 Chapter Summary

In this chapter, we briefly reviewed the major concepts in information theory in the context of brain dynamics. First, we explained the Shannon communication model. We interpreted the human brain as an information processing and communication system by establishing an analogy with Shannon's communication system. Then, we briefly described the concept of information entropy, which is the fundamental measure of Shannon information. We investigated how the human brain and mind can be related by using the concepts of entropy. We discussed the informative nature of entropy, which is associated with the concepts of complexity, randomness, disorder, and uncertainty.

We emphasized that Shannon entropy quantifies the amount of average information for specifying the state of the system or its level of uncertainty and unpredictability. We noted that the entropy of a system is high if it occupies many states in its state space with equal probability. This refers to maximum uncertainty, maximum information, and minimum knowledge. We also explained the relationship between the uncertainty and amount of the information: the more uncertain an event is the more information the describing message carries. Thus, a large number of bits are necessary to describe all the possible outcomes of the system.

In the last part, we described the rationale behind our information-theoretic approach for the analysis of neuroimaging data. We stated that entropy can be used as a powerful explanatory tool for investigating the neural processes which represent a cognitive state. We also discussed the advantages of using information-theoretic methods in

fMRI analysis. We went through some examples of information-theoretic studies that have been carried out in neuroimaging data analysis.

CHAPTER 4

A NEW INFORMATION THEORETIC METHOD FOR REPRESENTING MENTAL STATES OF COMPLEX PROBLEM SOLVING

> "The universe is very large, and its boundaries are not known very well, but it is still possible to define some kind of a radius to be associated with it."
>
> Richard P. Feynman

In this study, we propose an information theoretic method for the representation of brain activity during a complex problem solving task. Despite a considerable focus in the literature on complex problem solving, neural mechanisms underlying this process have not been studied in terms of the information content of the brain signals.

Complex problem solving is a fundamental ability of the human mind and it is considered as an essential characteristics of survival. It enables us to overcome the intricate issues we face, use the opportunities around us and gain control over the future. Thus, complex problem-solving skills are a critical part of life, both as individuals and as organizations.

At the World Economic Forum in 2015, complex problem solving was identified as one of the most important competencies required in the future [142]. Again, at the World Economic Forum in 2020, critical thinking and problem solving were at the top of the list of tomorrow's job skills [143]. Therefore, it is important to develop a full understanding of their cognitive and neural bases.

Complex problem solving interacts with many cognitive processes, which requires the distributed activity of numerous parts of the brain. Many neuroscientific studies on complex problem solving focused only on a set of predefined regions of the brain and ignored the rest. This may have led to some important findings being overlooked. It is important to use a holistic approach for the better characterization of the processes involved in complex problem solving and the underlying neural networks that support this task.

In this chapter, we start by briefly describing the complex problem solving task. Then, we describe the suggested computational model to explore the neural and mental underpinnings of this process.

4.1 Complex Problem Solving Process

Although several studies deal with the question of what complex problem solving is, there is no common agreement on the definition. According to Simon & Newell (1971), "the term problem solving refers to a higher-order cognitive function directed toward identifying problems with the current state and generating and implementing potential solutions to achieve a goal state" [10]. Frensch & Funke (1995) stated that "CPS takes place for reducing the barrier between an initial state and a goal state with the help of cognitive activities and behavior. Start state, goal state, and barriers determine the complexity and change dynamically over time. It demands the use of cognitive, emotional, and social abilities, and knowledge" [144].

Y. Wang et al. (2010) defined problem solving as "a high-order cognitive task of the human brain that requires various cognitive processes to work together, such as abstraction, searching, learning, decision making, inference, analysis, and synthesis based on internal knowledge representation" [11].

4.2 TOL Game as a Complex Problem Solving Experiment

Several cognitive tests have been developed to assess problem-solving ability of humans [145]. In the context of this study, we focus on a standard game, called Tower of London (TOL), which is frequently used in neuropsychology to measure planning and problem-solving abilities. The TOL game is accepted as an ideal example of a complex well-structured problem.

In the TOL game, there are three balls of different colors and three rods of different length. Each of the three rods of descending lengths could hold only 3, 2, or 1 ball, respectively. For each puzzle, two configurations are shown: an initial state and a goal state. In order to proceed from initial state to the goal state, the balls are moved one at a time from one rod to another. Subjects are asked to move from the initial state to the goal state using the minimum number of moves [146] (Figure 4.1). Problem complexity is related to the number of moves that depend on initial and goal positions. Participants are asked to plan mentally a sequence of moves, then execute their plan.

Complex problem solving involves two cognitive processing phases: planning and execution. These phases are similar to Simon and Newel's problem-solving model [10]. Their model proposed three main phases, namely, (1) the construction of a problem representation; (2) searching for the appropriate operators to solve the problem (elaboration or solution formulation); and (3) implementing the solution or execution. The planning phase incorporates the first two of these processes [147, 148].

Planning phase involves the construction of a problem representation and formulation of solution. In this phase, an efficient transformation with the minimum number of moves is predicted by looking mentally ahead. **Execution phase** involves the implementation of the solution plan, generated in the planning phase.

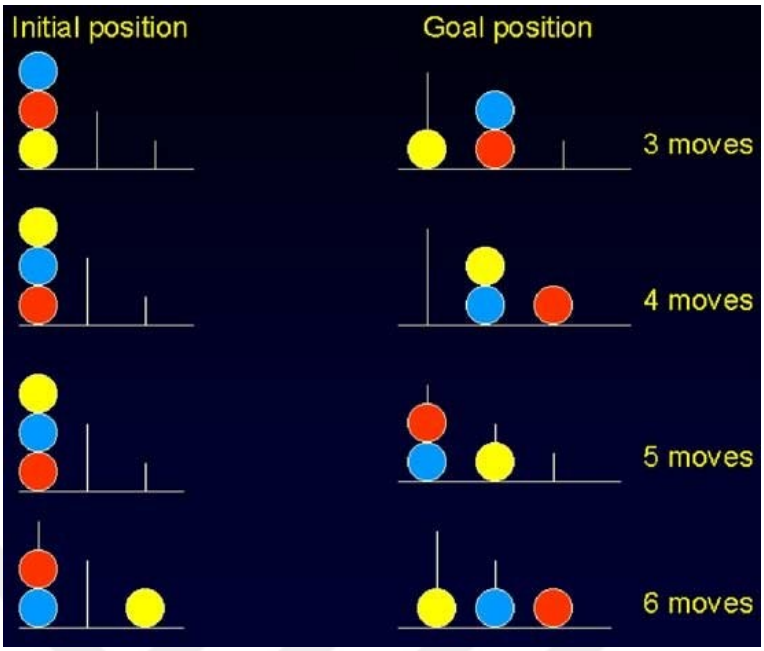


Figure 4.1: A set of Tower of London (TOL) problems that can be solved in three to six moves [6].

During planning, the problem solver creates a mental representation of the problem that is called the problem space [10]. This representation contains information about possible configurations that the problem may take, information about operators, knowledge from previous experiences with similar problems, and strategic knowledge [149]. Once the representation is formed, the solution formulation is made by exploring the problem space with an iterative process of creating and testing hypotheses. The effectiveness of the solution path depends on several factors, for instance, how extensively a problem solver explores the problem space to find optimum solution [147].

Planning involves several sub-processes such as strategy formation, coordination and sequencing of mental functions, and holding information online [150]. It incorporates a set of recursive processes, such as:

1. Decision-making to decide which ball to move and where to move.
2. Mental imagery to imagine the board in each step.
3. Working memory to encode and maintain the intermediate steps.

Execution only requires the retrieval of the steps from memory and the motor execution of those steps. Therefore, it is supposed that the planning phase demands more computational power compared to the execution phase [146].

Although the problem-solving models propose a sequential set of processes, the recent studies suggested that the interplay between these processes can be quite dy-

namic, depending on the demands of the problem [147]. Most of the neuroscientific studies in the literature have only focused on the planning phase since the execution phase might cover the re-planning process when the first plan is not successful to reach the goal state. In some cases, where subjects are unable to create a complete plan during planning, then online planning occurs during execution. In such cases, the brain regions responsible for generating the plan is expected to be activated during both the planning and implementation phases.

4.3 Information Content of Brain Regions

In this study, we investigate the information content of brain anatomical regions during the planning and execution phases of complex problem solving for strong and weak problem solvers, using first-order Shannon entropy.

Our major assumption is that the low entropy brain regions are more intimately involved in complex problem solving processes compared to the high entropy regions. Therefore, we investigate the relationship between the problem-solving task phases and the entropy measures. We measure the information content of brain regions by Shannon entropy at each anatomical region for two main phases of problem-solving and also for expert and novice players.

In Chapter-3, we mentioned that, Shannon entropy quantifies the average amount of information required to predict a pattern or a specific state of a stochastic system, or the outcome of a probabilistic process. Based on the Shannon entropy definition, we assume that measuring the information content of anatomical regions would allow us to make more accurate inferences about their states (as an outcome of a cognitive process).

As we reviewed earlier, there are various ways of estimating the entropy in human brain depending on the definition of the random process. In this study, we propose two new definitions of entropy to measure the information content of brain regions:

- **Static entropy**, measures the entropy of each anatomical region, independent of time. For each anatomical region, static entropy is defined over a period of time for the planning and execution phases separately. In this approach, we assume that each anatomical region is represented by average BOLD signals of voxels that reside in that region. Then, we assume that each planning or execution time instant of the average BOLD signal is a random variable. Static entropy is used to analyze the information content of the anatomical regions while the subjects perform the planning or execution phases of TOL problems.
- **Dynamic entropy** is defined as a function of time at each anatomical region. To estimate dynamic entropy, we take each voxel BOLD signal intensity in an anatomical region and assume that each voxel intensity value at each time instant is a random variable. Dynamic entropy is used to monitor the pattern of entropy fluctuations in anatomical regions and, accordingly, to analyze the

change of the information content of anatomical regions at each time instant during the resting state and the phases of the TOL problem-solving.

Using these definitions, we investigate the relationship between complex problem solving phases and entropy values of anatomical regions.

4.3.1 Estimating Static and Dynamic Entropy for Brain Regions

In this section, we describe our method of estimating the entropy of anatomical regions when solving complex problems. Since information and entropy definitions depend heavily on the definition of the random process and probability distribution, we briefly describe commonly used probability density estimation methods. Next, we explain our method of estimating probability distributions and then, estimating dynamic and static entropy.

Probability Density Estimation

Modeling a probability density from a random distribution is a well-studied problem in statistics, referred to as density estimation.

There are two approaches to density estimation:

- **Parametric approach** assumes that data are drawn from one of a known parametric family of distributions. A major limitation of this approach is that, the selected density might be a poor model of the distribution that produces the data, resulting in poor predictive performance [113].
- **Non-parametric approach** estimates the signal's probability distribution directly from its data points without assuming a particular form about the distribution of the data. In this model-free approach, data are allowed to speak for themselves in estimating the probability density more than they do if the probability density is restricted to fall in a particular parametric family.

Since the fMRI data would not fit the common analytical forms, we focus on non-parametric representation of the probability density function (pdf) to avoid making assumptions about the distribution of our data. The simplest form of non-parametric density estimation is the histogram.

Histogram

Consider a set of n data samples $x_1, ..., x_n$, as an outcome of a probabilistic event, represented by a random variable, x. In order to build a histogram, the sample data values are partitioned into a number of intervals (bins) of width Δ_i and counted the number n_i of observations of x falling in bin i. This count is then divided by the total

number N of observations and by the width Δi of the bins to find probability values for each bin, given by the equation below [113]:

$$p(x_i) = \frac{n_i}{N\Delta_i} \tag{4.1}$$

The final form of the density estimate depends on the bin width and starting position of the bins. The bins are typically chosen to have the same width.

Histograms can be useful for rapid analysis of data in one or two dimensions. Unfortunately, they are not suitable for most density estimation applications. They provide only a rough representation of the underlying distribution. Furthermore, the estimated density might have discontinuities, that are not due to the underlying distribution, but the bin edges. This makes difficult to understand the structure of the data. Another limitation of the histogram approach is the curse of dimensionality, since the number of bins grows exponentially with the number of dimensions [113].

There are two commonly used non-parametric approaches for density estimation, which have better scaling with dimensionality than the histogram model: kernel density estimators (KDE) and K-nearest-neighbours (kNN). In this study, we use the kernel density estimation approach. Therefore, we focus on this technique here. Note that, it can be shown that both kNN and KDE converge to the true probability density as total number of samples goes to infinity [151].

Kernel Density Estimation (KDE)

This technique is also called the Parzen-Rosenblatt window or **Parzen window** method [152, 153]. Similar to a histogram, the kernel density estimation builds a function to represent the probability distribution which is an approximation of the true density estimated using the sample data. The kernel density estimator generate the probability density by creating individual probability density curves for each data value, then summing the curves. The resulting probability density estimation is a smooth, continuous probability curve which is a sum of small density curves centered at each data point. Each data point makes an equal contribution to the final probability density.

Figure 4.2 shows the histogram (left) and the corresponding kernel probability density estimate (right) graphs plotted based on the same data values. Six kernels corresponding to six data points are indicated by red dashes. The blue line shows the resulting kernel density curve [7].

Kernel density estimation approximates the distribution of data by a mixture of continuous distributions K, called kernels, that are centered at x_i data points and have a scale (bandwidth) equal to h. Let, x_i is a random variable containing the data points

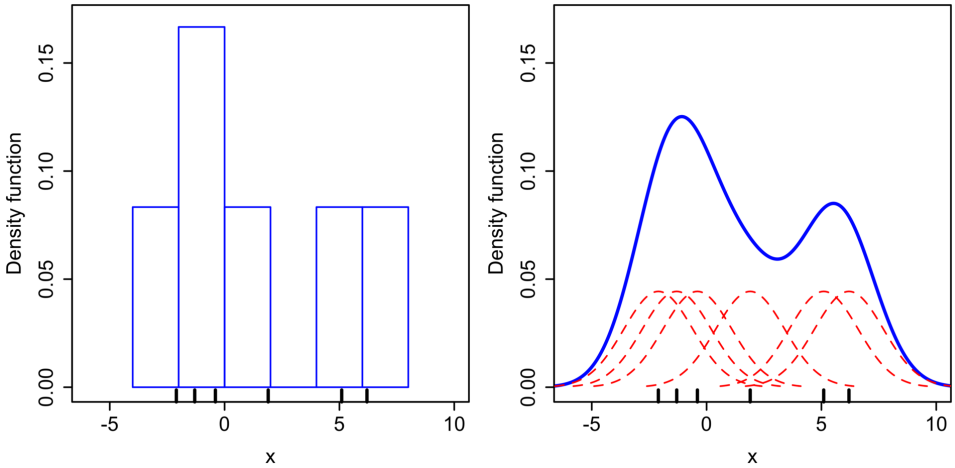


Figure 4.2: The histogram (left) and the corresponding kernel probability density estimate (right) graphs plotted based on the same data values. Six kernels corresponding to six data points are indicated by red dashes. The blue line shows the resulting kernel density curve [7].

$x_1, x_2, \ldots, x_N$. N is the number of sample data points. The kernel density (Parzen window) estimation equation which returns the probability density for x_i is:

$$p(x_i) = \frac{1}{Nh} \sum_{i=1}^{N} K\left(\frac{x - x_i}{h}\right). \tag{4.2}$$

Two parameters play a key role in kernel estimation: kernel function, K, and bandwidth, h.

Kernel function determines the shape of the curves. It is a mathematical function that returns a probability for a given data value. Kernel K integrates to unity. It also requires to be symmetric, so that $K(x) = K(-x)$. There are several types of kernel functions, such as Gaussian (normal), Epanechnikov (parabolic), uniform (rectangular), triangular, biweight (quartic), triweight, cosine, tricube, quadratic. The most popular and widely used kernel function is the Gaussian kernel.

The Gaussian kernel function K, with mean μ and standard deviation h is defined as,

$$K(x|\mu, h) = \frac{1}{h\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2h^2}} \tag{4.3}$$

The standard deviation, h, is sometimes called the *bandwidth*, which is a crucial parameter to be selected for the Kernel Density Estimation method. The bandwidth specifies the smoothness of the density. It determines the number of data points or

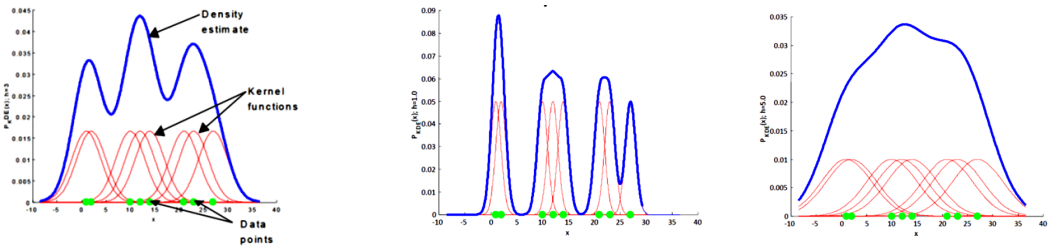


Figure 4.3: a) Kernel density estimation with the *optimal bandwidth* (left), b) a *small bandwidth* (middle), and c) a *large bandwidth* (right) [8].

window of data points from the sample set that contribute to estimating the probability for a given sample, x.

Choosing the proper bandwidth size, h, for the fMRI data controls the bias-variance trade-off: too small value may cause the estimator to show insignificant details. Thus, the result will not be general enough to correctly cover new examples. It yields a density estimation with many spikes at the observations. On the other hand, too large value implies high bias but small variance, resulting in a slowly varying curve. It leads to oversmoothing of the information contained in the data, which, in consequence, may mask some of the essential features. The probability distributions with optimal, small and large bandwidths are shown in Figure 4.3.

There are a number of suggested techniques for automatic (data-driven) bandwidth selection [154]. One of the most popular among these techniques is the Silverman bandwidth method [155], which is considered as a robust, sophisticated, and safe procedure in univariate distributions and is widely used in practice. It works fast and produces an optimum bandwidth value for the distributions close to normal. The Silverman's formula for the optimum bandwidth h is:

$$h = 0.9\, min(\sigma, \frac{IQR}{1.34})\, n^{-\frac{1}{5}} \tag{4.4}$$

where σ is the standard deviation of the sample data; n is the number of data points; IQR is the interquartile range which is the difference between the 75th percentile (Q3) and the 25th percentile (Q1) of the data values (i.e., $IQR = Q_3 - Q_1$). A *percentile* is a score at or below which a given percentage of scores in its frequency distribution falls.

4.3.1.1 Dynamic Entropy Estimation for Brain Regions Using Voxel Time Series

Dynamic entropy is used to analyze the change of the information content of brain regions at each time instant during the resting state and the complex problem solving (CPS) phases. In this approach, we assume that the entropy of a brain region changes

as a function in time. Then, the values of entropy over time can provide information about the CPS process and sub-tasks of CPS, namely planning and execution phases.

In order to estimate a dynamic entropy for each region, we use voxel intensity values, $v_i(t)$, measured at location (x_i, y_i, z_i) at time t. For each region and time instance, the probability distributions are estimated using intensity values of voxels residing in that region.

Formally speaking, let, r shows the index of a brain region, where $r = \{1, 2, 3, \ldots, R\}$, R is the total number of anatomical regions; $v_i(t)$ is i th coordinate voxel intensity value in the region r, at time t where $i = \{1, 2, \ldots, n\}$; n_r is the number of voxels in the region r; h_r represents the bandwidth for the region r during a session for a subject. For each time instance, recorded in an fMRI session, the probability density function of the voxels in a region $P_r(v(t))$ is estimated using the kernel probability density estimation method, from the following equation:

$$P_r(v(t)) = \frac{1}{n_r h_r} \sum_{i=1}^{n_r} K\left(\frac{v(t) - v_i(t)}{h_r}\right), \tag{4.5}$$

where K is the Gaussian kernel function, which is given in equation 4.3. We use the locally varying estimation of the bandwidth, h_r, for each region, based on the characteristics of voxel intensity values residing in that region. We apply Silverman's bandwidth or *Silverman's rule of thumb (Nrd0)* bandwidth method [155]. The formula for the optimum bandwidth h_r is:

$$h_r = 0.9\, min(\sigma(v(t)), \frac{IQR(v(t))}{1.34})\, n_r^{-\frac{1}{5}}, \tag{4.6}$$

where σ is the standard deviation of the voxel intensity values in the region r in time t; n_r is the number of voxels in the region r. IQR is the interquartile range which is the difference between the 75th percentile (Q3) and the 25th percentile (Q1) of the voxel values in the region r in time t (i.e., $IQR = Q_3 - Q_1$). An l-th percentile is a score at or below which a given percentage l of scores in its frequency distribution falls.

The examples of the resulting pdf's are given in Figure 4.4. The graphs show the probability density functions that are plotted for a brain region, using the planning time series of a session. The pdf's for each planning time instant are shown on the same graph.

For $\forall v(t) \in r$, the dynamic entropy of the brain region r at time t is estimated by,

$$H_r(v(t)) = -\sum_{\forall v(t) \in r} P_r(v(t))\, log_2 P_r(v(t)). \tag{4.7}$$

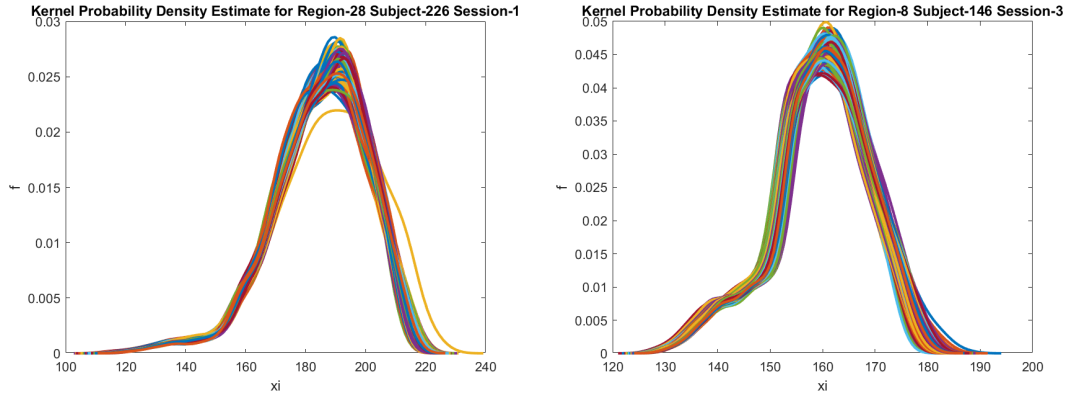


Figure 4.4: Probability density functions for Right Middle Frontal gyrus for subject-226 session-1 (Left), and Right Caudate for subject-146 session-3 (Right). The pdf's are plotted using the planning time series in the session. The pdf's for each planning time instant are shown on the same graph.

The algorithm for the estimation of dynamic entropy for an anatomical region in a time-instance is given in figure 4.5.

4.3.1.2 Static Entropy Estimation for Brain Regions

We estimate the static entropy using two different approaches. The first approach estimates the static entropy using representative time series of anatomical regions. We use this approach to obtain and compare the entropy values of the planning and execution phases for each anatomical region for successful and unsuccessful runs. In the second approach, we estimate static entropy using voxel time series residing in a single brain region. We use this estimates of pdf to measure and compare the entropy values of anatomical regions for successful and unsuccessful runs. We also compare the entropies of the planning and execution tasks.

4.3.1.2.1 Static Entropy Estimation Using Representative Time Series

In this approach, we assume that each brain anatomical region can be represented by a time series obtained by averaging all the voxel time series, which resides in that region. We define each time instance of the representative time series of an anatomical region as a random variable. We, then, estimate a probability distribution function for each anatomical region and each cognitive state, namely, planning and execution.

Let us now formally describe the static entropy estimation method for each brain region. Suppose that, fMRI BOLD signals recorded during a complex problem solving (CPS) session consists of time series $v_i(t)$ at each voxel coordinate i to represent the

Dynamic Entropy Estimation Algorithm

Input: Voxel time-series and labels for subjects and sessions
Output: Dynamic entropy values by region and time-instance

```
for each subject and session do;
    read session data and labels;
    for each region (roi) do;
        extract voxel time-series within that region;
        for each time-instance (i) do;
            % calculate optimum bandwidth bw-opt using equation 4.6
            bw_opt = 0.9*min(std(data(roi,i)),iqr(data(roi,i))/1.34)*length(data(roi,i))^-0.2;
            % calculate probability density estimate using equation 4.5
            pdf_est = ksdensity(data(roi,time-instance),'Bandwidth',bw_opt);
            % to handle 0 probabilities, adding 'eps' to all probabilities
            pdf_est ← pdf_est + eps;   (to handle 0 probabilities, adding 'eps' to all probabilities)
            % calculate Entropy -sum(p*log2(p)) using equation 4.7 for region (roi) and time-instance (i)
            log2Pi = log2(pdf_est);
            Pilog2Pi = pdf_est * log2Pi;   % calculate inner product
            Kernel_Entropy(roi,i) = - sum(Pilog2Pi);
        end;
    end;
end;
```

Figure 4.5: The algorithm for the estimation of dynamic entropy for an anatomical region in a time-instance.

neural activity of the underlying CPS task. The representative time series, $X_r(t)$ for an anatomical region r is estimated by,

$$X_r(t) = \frac{1}{n_r} \sum_{\forall v_i \in r} v_i(t), \tag{4.8}$$

where n_r is the number of voxels in region r.

In order to estimate the pdf of an anatomical region, we employ the Parzen window - kernel density estimation method. For each region r, we estimate two probability distribution functions: One for the planning phase and the other one for the execution phase, using the label of each time instance. The probability distribution functions in region r, for *planning* phase $P_{pr}(x)$ and *execution* phase $P_{er}(x)$, are estimated as follows:

$$P_{pr}(x) = \frac{1}{h_r n_p} \sum_{t=1}^{n_p} K\left(\frac{x - X_r(t)}{h_r}\right) \tag{4.9}$$

and

$$P_{er}(x) = \frac{1}{h_r n_e} \sum_{t=1}^{n_e} K\left(\frac{x - X_r(t)}{h_r}\right), \tag{4.10}$$

where n_p and n_e are the number of time instances for planning and execution phases respectively. X_r is the representative time series for a region r. K is the Gaussian kernel smoothing function, described in equation 4.3; and h_r is the bandwidth value. We implement the kernel density estimation through the Matlab *ksdensity* function. It returns the computed probability density estimate using the specified kernel function and the bandwidth of the kernel. In this approach, we use an empirically specified bandwidth value.

Figure 4.6 shows the estimated probability density functions for the Left Medial Orbitofrontal Gyrus for planning phase using two different bandwidth values. The red dashes corresponds to a small bandwidth, blue line corresponds to a large bandwidth. As it is illustrated in Figure 4.6, the estimated distribution is unimodal Gaussian density for a **single subject** and a Gaussian mixture for **multiple subjects**. The Gaussian mixture of multiple subjects can be attributed to the substantial differences among the subjects. Similar behaviour is observed for all of the anatomical regions.

Once we estimate the probability distribution functions in region r, for *planning* phase $P_{pr}(x)$ and *execution* phase $P_{er}(x)$, the static entropy of the anatomical region r for *planning* phase is then estimated by,

$$H(P_{pr}(x)) = -\sum_{\forall x \in r} P_{pr}(x)\, log_2\, P_{pr}(x). \tag{4.11}$$

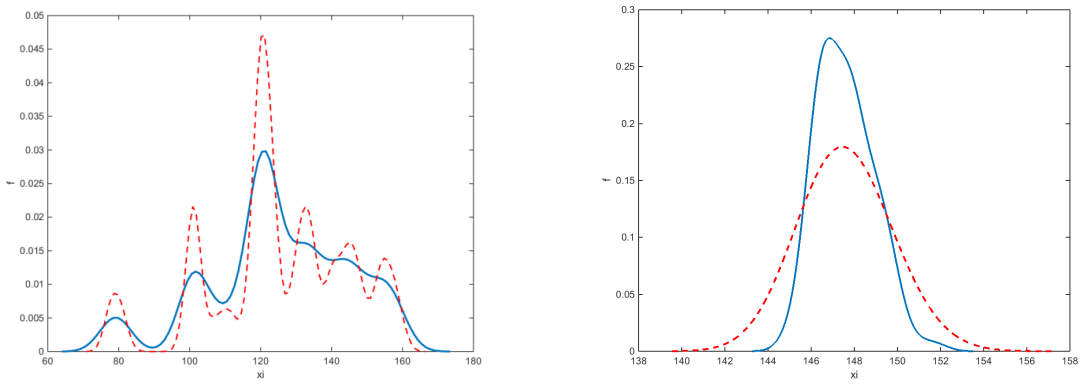


Figure 4.6: Probability density functions using two different bandwidths when the representative time series are used as random variable, for *Left Medial Orbitofrontal Cortex* across all subjects (Left) versus single subject (Right). The red dashes corresponds to a small bandwidth, blue line corresponds to a large bandwidth. In each plot Gaussian Kernels are used. Note that across subject density is a Gaussian mixture whereas the single subject density is a unimodal Gaussian function.

And the static entropy of the anatomical region r for *execution* phase is then estimated by,

$$H(P_{er}(x)) = -\sum_{\forall x \in r} P_{er}(x)\, log_2\, P_{er}(x). \tag{4.12}$$

Note that, in the above formulations, representative fMRI time series $X_r(t)$, of each anatomical region, for all sessions of a subject were concatenated, separately for *planning and execution* sections. Then, the probability distributions were calculated for *planning and execution* phases for each subject and anatomical region.

Figure 4.7 shows the algorithm that summarizes the processes for the estimation of static entropy of the anatomical regions using representative time series for the planning task of a subject.

4.3.1.2.2 Static Entropy Estimation Using Voxel Time Series

To compare the overall entropies of anatomic regions for successful and unsuccessful subjects-sessions, we estimate the static entropy using voxel time series residing in the underlying region. In this approach, we first estimate the dynamic entropies for each region and each planning and execution time instances, using the distribution of voxel intensity values residing in that region. Then, we separate the successful and unsuccessful runs. We estimate average entropy values by anatomic regions using the entropy values for planning and execution time instances.

Static Entropy Estimation using Representative Time-series

```
Input: Voxel time-series and labels for subjects and sessions
Output: Static entropy values by region
for each subject do;
    for each session do;
        read session data and labels;
        for each time-instance do;
            for each region do;
                average voxel time-series within that region using equation 4.8;
            end;
        end;
        take planning and execution time-series for the session;
        save the corresponding region averaged data and the labels;
    end;
    read region averaged data for each sessions of the subject, and corresponding labels;
    concatenate all session's labels into one label file that includes pl and ex labels;
    concatenate all session's data into one data file that includes pl and ex region averaged values;
    for each region do;
        read region averaged values of all the planning time-series for the subject;
        calculate optimum bandwidth for the planning time-series of the region using equation 4.6;
        calculate probability density estimate for the planning time-series using equation 4.9;
        calculate Entropy of the region using equation 4.11 for the planning phase;
    end;
end;
```

Figure 4.7: The algorithm that summarizes the processes for the estimation of static entropy of the anatomical regions using representative time series for the planning task of a subject.

Let, r shows the index of a brain region, where $r = \{1, 2, 3, \ldots, R\}$, R is the total number of anatomical regions; $v_i(t)$ is i th coordinate voxel intensity value at planning or execution time t_{pe}, where $i = \{1, 2, \ldots, n\}$; n_r is the number of voxels in the region r; h_r represents the bandwidth for the region r during a session for a subject, estimated by equation 4.6. K is the Gaussian kernel function. For each planning and execution time instance, the probability density function of the voxels in a region $P_r(v(t))$ is estimated using kernel probability density estimation, from the following equation:

$$P_r(v(t)) = \frac{1}{n_r h_r} \sum_{i=1}^{n_r} K\left(\frac{v(t_{pe}) - v_i(t_{pe})}{h_r}\right). \tag{4.13}$$

For $\forall v(t_{pe}) \in r$, the entropy of the brain region r at time t_{pe} is estimated by,

$$H_r(v(t)) = -\sum P_r(v(t)\ log_2\ P_r(v(t)). \tag{4.14}$$

Then, for each *successful run* s, we calculate the average entropy values by region, using planning and execution time instances entropy values, by the equation below

$$H_{sr} = \frac{1}{n_{pe}} \sum_{\forall t_{pe}} H_r(v(t)), \tag{4.15}$$

where n_{pe} is the number of time instances for planning and execution phases. For each *unsuccessful run* u, the average entropy values by region, is obtained using planning and execution time instances entropy values,

$$H_{ur} = \frac{1}{n_{pe}} \sum_{\forall t_{pe}} H_r(v(t)). \tag{4.16}$$

Figure 4.8 shows the algorithm that summarizes the processes for the estimation of static entropies of the anatomical regions by task, using voxel time series for the successful runs

4.4 Estimating Static and Dynamic Brain Networks

Representing the fMRI signals by brain networks is very crucial to understand the information exchange among the anatomic regions during the underlying cognitive states. The complexity of human behavior is based on the organization of neurons as anatomical circuits with precise functions. Each neuron is a node of one or more networks, which carries information about one or more specific cognitive functions.

Static Entropy Estimation using Voxel Time-series

```
Input: Entropy values by region & time-instance, and labels for successful sessions
Output: Static entropy values by region & task (game, pl or ex) for successful sessions
for each successful session do;
  read session data and labels;
  for each region do;
    average the entropy values for planning and execution time instances using equation 4.15
    average the entropy values for planning time-instances using equation 4.15
    average the entropy values for execution time-instances 4.15
  end;
  save the averaged entropy values by region for TOL task for successful sessions
  save the averaged entropy values by region for planning task for successful sessions
  save the averaged entropy values by region for execution task for successful sessions
end;
```

Figure 4.8: The algorithm that summarizes the processes for the estimation of static entropies of the anatomical regions by task, using voxel time series for the successful runs.

One of the brain's basic organizational principles is that nerve cells with similar characteristics can produce different actions depending on how they are connected with each other. The specific patterns of neural connections and how these circuits are functionally organized in different brain regions drives our thoughts and behavior.

Recent methods highlight the importance of computational techniques, specifically brain network models, for understanding the brain and designing the brain-computer interfaces. Brain networks enable us to analyze the interactions across the anatomical regions and model the complex cognitive processes, such as, complex problem solving, emotions and memories. In this study, we propose a novel method to estimate static and dynamic brain networks, which represent the planning and execution phases of CPS from the recorded fMRI data. The proposed network models enable us to investigate the complex problem solving phases, which lead to activities among brain regions. We also examine various functional connectivity patterns.

Our dynamic and static entropy findings led us to model the functional connectivity of brain regions by **Kullback-Leibler (K-L) divergence**, which is also called **Relative Entropy**. It measures the similarity or the statistical difference between two probability distributions. Kullback–Leibler divergence was introduced by Solomon Kullback and Richard Leibler in 1951 as the directed divergence between two distributions [156]. The K-L divergence of two random variables is the expected value of the difference of the self-information between two random variables. The value of

K-L divergence depends on the selection of the random variable, which we take the expectation.

Our major assumption is that *the degree of co-activation between two brain regions can be measured by Kullback-Leibler divergence*. Therefore, the measure of K-L divergence between the anatomical regions can be used as the **arc weights** of the brain network formed among the anatomical regions.

Based upon this assumption we estimate two types of brain network, namely, static and dynamic brain networks. First, **dynamic brain networks** are estimated from the distribution of voxel time-series for each anatomical region, at each time instant. Kullback-Leibler divergence between the anatomical regions are, then, estimated for the distribution functions of the region pairs. We represent the planning and execution phases of each strong and weak problem solvers by the estimated node adjacency matrix of the brain network.

Second, **static brain networks** are estimated to investigate the behaviour of the planning and execution tasks. Static brain networks are estimated by taking the expected value over all time instances to represent the K-L divergences for all the planning and execution phases.

4.4.1 Kullback-Leibler Divergence Estimation

Consider two probability distribution functions, $P_k(x)$ and $P_l(x)$, defined on the same probability space, over a discrete random variable, $x \in \mathcal{X}$. Kullback–Leibler divergence between $P_k(x)$ and $P_l(x)$ is defined [156] as,

$$D_{\mathrm{KL}}(P_k \parallel P_l) = -\sum_{x \in \mathcal{X}} P_k(x) \log \left(\frac{P_k(x)}{P_l(x)} \right). \tag{4.17}$$

Equation (4.17) shows the expectation of the logarithmic difference between the probability distribution functions, P_k and P_l, where the expectation is taken over the probability distribution function, P_k.

In Bayesian inference language, $D_{\mathrm{KL}}(P_k \parallel P_l)$ measures the amount of information gained when we revise our beliefs from the pdf, P_l to the pdf, P_k. In the context of this study, we estimate the probability distribution function P_r of each brain region r and measure the Kullback-Leibler divergence between two brain regions. Then, we construct a brain network across the anatomical regions, where the edge weights correspond to the estimated K-L divergences; assuming that the degree of co-activation between two anatomical regions is measured by the K-L divergences.

In this study, we estimate two types Kullback-Leibler divergences, depending on the random variable defined for each anatomical region. The first one is called static Kullback-Leibler divergence, which is estimated from the probability distribution functions of each brain regions, for the entire time series of planning and execution

phases. The second one, called dynamic Kullback-Leibler divergence, is estimated for each time instance across the brain regions.

The following subsections introduces the suggested estimation methods for Kullback-Leibler divergences.

4.4.1.1 Estimation of Static Kullback-Leibler Divergence

In order to estimate a static brain network for planning and execution, we assume that each brain region can be represented by a time series obtained by averaging all the voxel time series, which reside in that region. We define each time instance of the representative time series of an anatomical region as a random variable. We estimate a probability distribution function for each brain region and each cognitive state, namely, *planning and execution.* As a result, we obtain a probability distribution function for each cognitive state and for each brain region.

Let us now formally describe the method for estimation of static Kullback-Leibler divergence to represent planning and execution phase across the anatomical regions. Suppose that, fMRI signals recorded during a Complex Problem Solving (CPS) session, consists of time series $v_i(t)$ at each voxel coordinate i to represent the neural activity of the underlying CPS task. The representative time series, $X_r(t)$ for an anatomical region r is estimated by,

$$X_r(t) = \frac{1}{n_r} \sum_{\forall v_i \in r} v_i(t), \tag{4.18}$$

where n_r is the number of voxels in region r.

For each region r, we estimate two probability distribution functions: One for the planning phase and the other one for the execution phase, using the label of each time instance. The probability distribution functions in region r, for *planning* phase $P_{pr}(x)$ and *execution* phase $P_{er}(x)$, are estimated by kernel probability density estimation method, as follows:

$$P_{pr}(x) = \frac{1}{h_r n_p} \sum_{t=1}^{n_p} K\left(\frac{x - X_r(t)}{h_r}\right) \tag{4.19}$$

and

$$P_{er}(x) = \frac{1}{h_r n_e} \sum_{t=1}^{n_e} K\left(\frac{x - X_r(t)}{h_r}\right), \tag{4.20}$$

where n_p and n_e are the number of time instances for planning and execution phases, respectively. K is the Gaussian kernel smoothing function, and h_r is the bandwidth of the Gaussian kernel for region r.

When we consider the time instances of a representative time series with *planning* label, the corresponding K-L distance between the probability distributions of two brain regions k and l, for $\forall\ l \neq k$ is, then, estimated by

$$D_{\mathrm{KL}}(P_{pk}(x) \parallel P_{pl}(x)) = -\sum_{x \in \mathcal{X}} P_{pk}(x) \log \left(\frac{P_{pk}(x)}{P_{pl}(x)} \right). \tag{4.21}$$

On the other hand, when we consider only the time instances with execution label, K-L divergence is estimated by

$$D_{\mathrm{KL}}(P_{ek}(x) \parallel P_{el}(x)) = -\sum_{x \in \mathcal{X}} P_{ek}(x) \log \left(\frac{P_{ek}(x)}{P_{el}(x)} \right). \tag{4.22}$$

Note that, in the above formulations, representative fMRI time series $X_r(t)$, of each brain anatomical region, for all sessions are concatenated, separately for *planning and execution* sections. Then, the probability distributions are calculated for *planning and execution* phases for each subject and brain region.

Figure 4.9 shows the algorithm for estimation of static Kullback-Leibler divergences to generate the directed static brain networks for the planning task of a subject.

As a baseline method, we estimate the Pearson correlation coefficients between the representative time series of each region, and generate a static network using them as arc weights, as will be explained in the experiment section.

4.4.1.2 Estimation of Dynamic Kullback-Leibler Divergence

In order to estimate a dynamic brain network, we directly use voxel intensity values, $v_i(t)$, measured at location (x_i, y_i, z_i) at time instance t. For each region and time instance, the probability distributions are estimated using intensity values of voxels residing in that region.

Formally speaking, let r be the index of an anatomical region, where $r = \{1, 2, 3, \ldots, R\}$ and R is the total number of anatomical regions; $v_i(t)$ is ith coordinate voxel intensity value at time t where $i = \{1, 2, \ldots, n\}$; n_r is the number of voxels in the region r; h_r represents the selected bandwidth for the region r during a session for a subject. For each time instance, recorded in an fMRI session, the probability distribution function over all voxels in a region $P_r(v(t))$ is estimated using kernel probability distribution function estimation, from the following equation:

$$P_r(v(t)) = \frac{1}{n_r h_r} \sum_{i=1}^{n_r} K \left(\frac{v(t) - v_i(t)}{h_r} \right). \tag{4.23}$$

Static Kullback-Leibler Estimation Algorithm

```
Input: Voxel time-series and labels for subjects and sessions
Output:  Arc-weight matrices for the planning task

for each subject do;
    for each session do;
        Read  voxel time-series and labels ;
        for each time-instance do;
            for each region do;
                average voxel time-series within the  region using equation 4.18;
            end;
        end;
        take planning and execution time-series for the session;
        save the corresponding region averaged data and the labels;
    end;
    Read region averaged data for each sessions of the subject, and corresponding labels;
    Concatenate all session's labels into one label file that includes pl and ex labels;
    Concatenate all session's data into one data file that includes pl and ex region averaged values;
    for each region (roi-1) do;
        Read region averaged values of all the planning time-series for the subject;
        Calculate optimum bandwidth for the planning time-series of the region using equation 4.6;
        Calculate probability density estimate for the planning time-series using equation 4.19;
        for each region (roi-2) do;
            if roi-2 = roi-1 then
                arc-weight matrice(roi-1, roi-2) ← 1
            Calculate optimum bandwidth bw-2 using equation 4.6;
            Calculate probability density estimate using equations 4.19;
            Normalize pdf_est1 and pdf_est2 by dividing each value into the sum of all the values.
            To handle zero probabilities, add 'eps' to all probabilities
            Calculate K-L distance from roi-1 to roi-2 using equation 4.21;
            if K-L distance roi1-to-roi2 < 0 then error;
            Calculate K-L distance from roi-2 to roi-1 using equation 4.21;
            if K-L distance roi2-to-roi1 < 0 then error;
            Fill the directed arc weight matrice using K-L distance values;
        end;
    end;
end;
```

Figure 4.9: The algorithm for estimation of Kullback-Leibler divergences to generate the directed static brain networks for the planning task of a subject.

The K-L distance between two anatomical regions k and l for $\forall\ k \neq l$ at time t is estimated by

$$D_{\mathrm{KL}}(P_k(v(t)) \parallel P_l(v(t))) = -\sum_{v(t)\in\mathcal{X}} P_k(v(t))\log\left(\frac{P_k(v(t))}{P_l(v(t))}\right). \tag{4.24}$$

Note that, in the above formulation, Kullback-Leibler divergences, estimated across each pair of brain region are considered as the arc weights of a brain graph at each time instant. Since we obtain a probability distribution function for each time instant, over all the voxels residing in an anatomical region, it is assumed that this approach gives a dynamic nature of brain connectivity between anatomical regions. Then, a connectivity graph is produced at each time instant, using this approach.

The algorithm for estimation of Kullback-Leibler divergences to generate arc-weights for the directed and undirected dynamic brain networks is given in figure 4.10.

4.5 Chapter Summary

In this chapter, we introduced our computational model, which provides the information-theoretic representation of brain activity during the complex problem solving task, using fMRI data.

First, we have defined the complex problem solving process and its main phases, namely planning and execution. Next, we describe our computational model, which proposes entropy measurement for identification of active regions and reveals brain states associated with complex problem solving. We developed two new definitions of entropy: static entropy and dynamic entropy. Static entropy is defined over a period of time, while dynamic entropy measures brain activity at each time instant.

We also described the Kullback-Leibler Divergence (relative-entropy) for defining the static and dynamic brain networks activated during the planning and execution phases of the complex problem solving.

Dynamic Kullback-Leibler Estimation Algorithm

```
Input: Voxel time-series and labels for subjects and sessions
Output: Arc-weight matrices for each time-instance
for each subject and session do;
    read session data and labels;
    for each  Pl or Ex time-instance (i) do;
        arc-weight matrice ← { } ;
        for each region (roi-1) do;
            % Calculate optimum bandwidth bw-1 using equation 4.6
            optH1 = 0.9*min(std(data(roi-1,i)),iqr(data(roi-1,i))/1.34)*length(data(roi-1,i))^-0.2;
            % Calculate probability density estimate using equations 4.23
            pdf_est1 = ksdensity(data(roi-1,time-instance),'Bandwidth',bw-1);
            for each region (roi-2) do;
                if roi-2 = roi-1 then
                    arc-weight matrice(roi-1, roi-2) ← 1
                % Calculate optimum bandwidth bw-2 using equation 4.6
                optH2 = 0.9*min(std(data(roi-2,i)),iqr(data(roi-2,i))/1.34)*length(data(roi-2,i))^-0.2;
                % Calculate probability density estimate using equations 4.23
                pdf_est2 = ksdensity(data(roi-2,time-instance),'Bandwidth',bw-2);
                % Normalize pdf_est1 and pdf_est2 by dividing each value into the sum of all the values.
                pdf_est1 = pdf_est1 /repmat(sum(pdf_est1,2),[1 size(pdf_est1,2)]);
                pdf_est2 = pdf_est2 /repmat(sum(pdf_est2,2),[1 size(pdf_est2,2)]);
                % to handle 0 probabilities, adding 'eps' to all probabilities
                P  =  pdf_est1 + eps;
                Q  =  pdf_est2 + eps;
                % Calculate K-L distance from roi-1 to roi-2 using equation 4.24
                    KL1  =   P * log2 (P/Q);
                    % resolving the case when P(i)=0
                    KL1(isnan(KL1)) ← 0;
                    % K-L distance from roi-1 to roi-2
                    KLdistance-r1r2 ← sum(KL1,2);
                    if KLdistance- r1r2 < 0 then error;
                % Calculate K-L distance from roi-2 to roi-1 using equation 4.24
                    KL2  =   Q * log2 (Q/P);
                    % resolving the case when P(i)=0
                    KL2(isnan(KL2)) ← 0;
                    % K-L distance from roi-2 to roi-1
                    KLdistance-r2r1 ← sum(KL2,2);
                    if KLdistance- r2r1 < 0 then error;
                % fill the directed arc weight matrice for the pl or ex time-instance:
                    arc_weight_matrice_directed (roi1, roi2) = KLdistance- r1r2;
                    arc_weight_matrice_directed (roi2, roi1) = KLdistance- r2r1;
                % fill the undirected arc weight matrice for the pl or ex time-instance:
                    avg_distance =( KLdistance- r1r2+ KLdistance- r2r1)/2;
                    arc_weight_matrice_undirected (roi1, roi2) = avg_distance;
                    arc_weight_matrice_undirected (roi2, roi1) = avg_distance;
            end;
        end;
    end;
end;
```

Figure 4.10: The algorithm for estimation of Kullback-Leibler divergences to generate arc-weights for the directed and undirected dynamic brain networks.

CHAPTER 5

COMPLEX PROBLEM SOLVING EXPERIMENT AND DATA ANALYSIS

> The enchanting charms of this sublime science reveal only to those who have the courage to go deeply into it.
>
> Carl Friedrich Gauss

In this chapter, we present the main findings of the data analysis, based on the proposed method described in the Chapter 4. We first explain the Tower of London (TOL) experimental setup, the fMRI data collection environment and the fMRI data properties. Then, we investigate the behavior of static and dynamic entropy variations across subjects, the main phases of CPS, and the anatomical regions.

In this context, we examine dynamic entropy variations of the complex problem solving task compared to resting states. Then, we analyze the entropy of anatomical regions during the planning and execution phases of complex problem solving. We reveal anatomical regions with low entropy, which we suggest to be regions activated by this task. We also compare the low entropy regions for strong and weak problem solvers. We discuss the differences in the ability to solve complex problems.

In the next part, we present the results of the static and dynamic brain network analysis. Using the Kullback-Liebler Divergence method, we estimate planning and execution phases prominent connections for successful and unsuccessful runs. We test the validity of the suggested brain networks by training a classifier with the arc-weights of the estimated dynamic brain networks.Finally, we measure the test performances of the planning and execution tasks.

5.1 TOL Experimental Setup and Nature of fMRI data

In this section, we explain our experimental setup, the fMRI data collection technique and the underlying structure of the data.

The fMRI data were recorded while the subjects played the Tower of London (TOL) game. 18 healthy subjects between the ages of 19 and 38 solved a computerized version of the TOL game [157]. Two configurations were shown for each puzzle, the

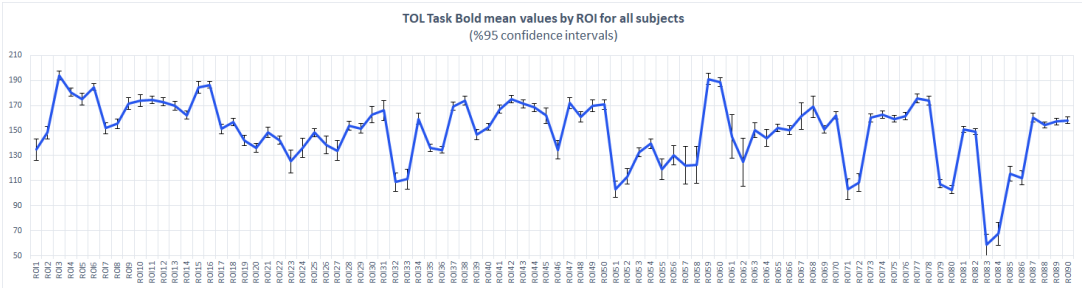


Figure 5.1: Averages of BOLD values across all sessions, based on anatomical regions with a 95% confidence interval.

initial state and the target state. The subjects were instructed to transform the initial state into the goal state using the minimum number of moves [148]. In each puzzle, it was required at least five or six moves to reach the target configuration. Problem complexity is related to the number of moves that depend on start and goal positions. The participants were directed to generate a solution plan before making their first move.

The experimental setup consists of two phases: In the first phase, called *planning*, the subjects are asked to make a plan to solve the problem. In the second phase, called *execution*, they play the game by moving a cursor. The planning phase is defined from the onset of the display of the puzzle until the first move of execution; the execution phase is defined from the first move until the end of the puzzle [148].

For each subject, 4 sessions were performed; each session contained 18 puzzles, where each puzzle is allocated a 15 seconds time limit. The first 5 seconds were allocated for planning only. Subjects could continue planning after the 5 seconds if they had not completed their solution plan [157]. The fMRI data has 590 time instances for each session. It contains TOL problem-solving data for a total of 72 sessions which has 1.296 puzzles in total. The average of all sessions BOLD values based on anatomical regions are shown in figure 5.1 with a 95% confidence interval.

The fMRI data is based on the 116 anatomical regions of interest (ROI) defined by the Anatomical Automatic Labeling (AAL) brain atlas [158] provided by the Montreal Neurological Institute (MNI). In this study, Cerebellum and Vermis regions were discarded, and the rest 90 regions were analyzed. The coordinates of voxels and the regions' size are the same for all subjects.

"The fMRI images were collected using a 3 T Siemens TRIO scanner with an 8-channel radiofrequency coil located in the Imaging Research Facility at Indiana University. The images were acquired in 18 5 mm thick oblique axial slices using the following set of parameters: TR=1000 ms, TE=25 ms, flip angle=60, voxel size=3.125 mm x 3.125 mm x 5 mm with a 1 mm gap. Preliminary data analysis was performed on the recorded fMRI data using statistical parametric mapping (SPM2 from the Wellcome Department of Cognitive Neurology, London) that included: image correction for slice acquisition timing (resampled to 2×2×2 mm voxels), spatial smoothing (with

a Gaussian filter of 8 mm; and high-pass filtering with 1/128 Hz cutoff frequency to remove low-frequency signals), motion correction" [146].

fMRI recordings are constructed as four-dimensional data in space-time. The intensity values, which indicate the degree of activity at each voxel coordinate, $v_i(t) = (x_i, y_i, z_i)$ are measured at each time instant, t. As a result, the data consists of time-series, $v_i(t)$, recorded at each coordinate (x_i, y_i, z_i) of a brain volume. In other words, at each time instant fMRI method generates a brain volume of intensity values, with a corresponding cognitive state label, namely *planning or execution* [159].

5.2 Entropy of Brain Anatomical Regions

In this study, we investigate the entropy of anatomical regions in the complex problem solving (CPS) process. In the following sections, we present analyses to examine the behavior of static and dynamic entropy variations across subjects, subtasks of CPS, and anatomical regions.

5.2.1 Dynamic Entropy Variations of Complex Problem Solving and Resting States

In the first set of experiments, we investigate the dynamic entropy variations for brain regions during the resting state and the CPS task and observe high and low entropy regions. Recall that low entropy indicates organized behaviors of the relevant region, while high entropy regions exhibit relatively random behaviors.

During the experiments, we notice that some subjects have relatively random fluctuations of dynamic entropy, compared to some other subjects. We assume that the fluctuations of entropy, which does not fit the event related stimuli, is an indicator of the level of expertise of a particular subject. The high fluctuations can also be interpreted as a relatively low focus of a subject during the complex problem solving task. We divided players into two groups based on the number of puzzles they successfully completed in a session and how many extra moves they made compared to the shortest path to complete a puzzle. We considered 12 subjects who successfully completed 756 puzzles in a total of 42 successful sessions as expert subjects and the rest as novice players.

In order to compare the behavior of the dynamic entropy measure in the anatomical regions, we estimate the dynamic entropy for each subject, for each session, for each time instance, and each anatomical region. The subjects are restricted to the 12 most successful subjects. Then, we average all the dynamic entropy measures across 42 sessions of 12 subjects, over an anatomical region and for each time instance. The comparison of dynamic entropy variations between the successful and unsuccessful players will be provided in the next section, which clarifies the reason why we select only the successful players in this set of experiments.

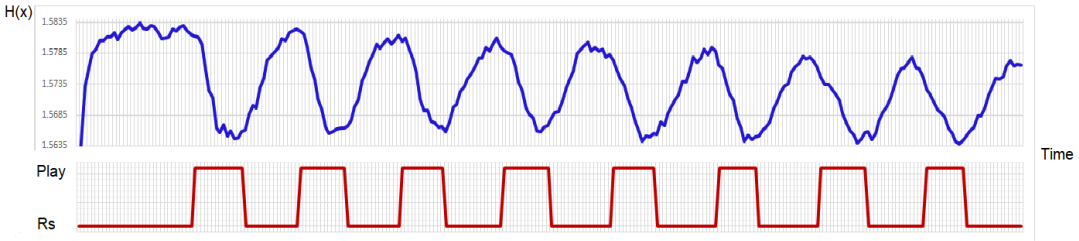


Figure 5.2: Dynamic Entropy vs. time for a minimum entropy region: Entropy fluctuations (blue) for **Left Superior Parietal** during TOL puzzle solving (averaged for 12 subjects, 42 sessions). *Play* shows the time instances when the subject plays the TOL game and *Rs* shows, when the subject is back to the resting state, in red plot. The first puzzle begins at 33. time-instance

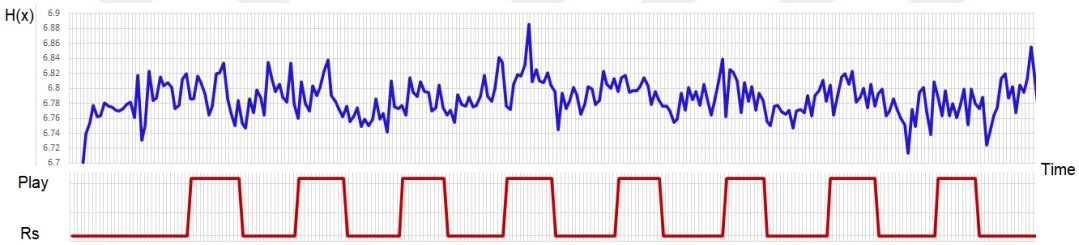


Figure 5.3: Dynamic Entropy vs. time for a maximum entropy region: Entropy fluctuations (blue) for **Right Insula** during a TOL session (averaged for 12 subjects, 42 sessions). *Play* shows the time instances when the subject plays the TOL game and *Rs* shows, when the subject is back to the resting state, in red plot.

For successful players, the measures of dynamic entropy is quite different in high and low entropy regions. We observe that dynamic entropy variations with respect to time is rather more structured for low-entropy anatomical regions, whereas it randomly fluctuates for high-entropy anatomical regions. This fact is exemplified in Figure 5.2 and Figure 5.3. While analyzing these figures, we should consider a slight shift in the entropy plot (blue) compared to the experimental setup plot (red), due to the delay of the hemodynamic response to the stimuli. Therefore, a slight shift to the right in the experimental setup plot would provide a more accurate match with entropy curve. This delay is to be taken into account while analyzing the entropy variations, in the rest of the dynamic entropy plots, provided in the subsequent sections [160].

As can be seen from Figure 5.2, the dynamic entropy in a low entropy region is highly correlated with the experimental setup. It is always low when the subject plays the TOL game and it is always high when the subject is in a resting state. This indicates that playing the TOL game generates a more organized signal in a low entropy region. In other words, an anatomical region that produces a more orderly outcome has a lower entropy value. "Order" indicates a decrease in the number of states within the

system, which makes the system more certain. Thus, the lower entropy value implies greater predictability of state of the system.

However, when we investigate dynamic entropy variations in the high entropy region, we observe that dynamic entropy shows an irregular course, appearing to be highly contaminated by noise. As it is seen in Figure 5.3, the dynamic entropy measure is not correlated with the experimental setup. As we have mentioned before, a high entropy value indicates greater unpredictability of the system behavior or outcome.

5.2.2 Dynamic Entropy Variations of Resting State, Planning and Execution Subtasks for Expert and Novice Players

In the next set of experiments, we investigate the behavior of dynamic entropy changes for the resting state fMRI signal and planning and execution phases of complex problem solving.

After grouping the subjects into expert and novice classes, we select an arbitrary expert player (*Subject-175*) and novice player (*Subject-881*), then, we estimate the dynamic entropy measures of each anatomical region of each subject. The minimum entropy for the expert player is observed in *Right Precuneus* region, and a minimum entropy measure is observed in *Right Lingual* region for novice player. Then, we investigated the behaviour of dynamic entropy in these most informative anatomical regions of expert and novice players.

For the **expert player**, the dynamic entropy in the resting state, planning-execution phases are highly correlated with the experimental setup (Figure 5.4 and Figure 5.6). In other words, the high entropy is observed during the resting state. As it is known, the entropy of a system is high if it occupies many states in its state space with equal probability. When the subject starts playing, the entropy is decreased in an anatomical region with the lowest information content. Thus, the predictability of its outcome (or state) is increased. As mentioned above, due to a slight delay of the hemodynamic response to the stimuli, one should consider a slight shift to the right in the red plot.

For the **novice players**, the dynamic entropy and task relations are relatively more random (Figure 5.5 and Figure 5.7), compared to that of the expert players. For the novice players, the dynamic entropy fluctuations are rather arbitrary during the resting state and while playing the TOL game, in an anatomical region with the lowest entropy.

5.2.2.1 Comparison of BOLD and Dynamic Entropy Variations of Resting State, Planning and Execution Subtasks

The above observations reveal that the Shannon entropy measure is more informative than the BOLD response to represent the neural activity related to the problem-solving process. In order to test this hypothesis, we compared the BOLD and entropy

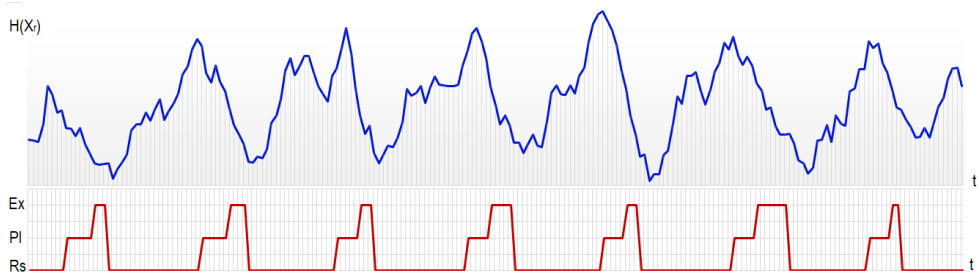


Figure 5.4: Expert Player: Dynamic Entropy fluctuations of a low entropy region, Right Precuneus during session-3 of an expert player, Subject-175 (blue). *Pl* shows the time instances when the subject is doing Planning, *Ex* shows the execution and *Rs* shows, when the subject is back to the resting state, in the red plot.

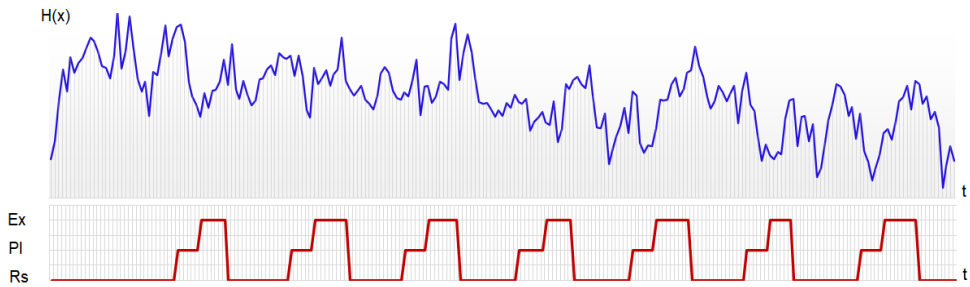


Figure 5.5: Novice player: Dynamic Entropy fluctuations for a low entropy region, Right Precuneus during first session for a novice player, Subject-881 (blue). *Pl* shows the time instances when the subject is doing Planning, *Ex* shows the execution and *Rs* shows, when the subject is back to the resting state, in the red plot.

values for several activated regions using the TOL problem-solving time-series. We concluded that the entropy measurement more precisely reflects the neural activity of anatomical regions during problem-solving than the BOLD measurement.

Figure 5.8 and Figure 5.9 show BOLD and entropy variations in an active anatomical region, respectively, during the TOL problem-solving session for an expert player. As shown in the graph, the Entropy values are compatible with the experimental setup; it falls rapidly when planning begins and rises sharply when execution is complete. However, in the case of BOLD values, the beginning and end of a puzzle are not as clearly reflected as that of entropy. It should be noted that BOLD values increase during neural activation, whereas the entropy values decrease with neural activation.

5.2.3 Analysis of Low Entropy Regions in Complex Problem Solving

The major assumption of this study is that the lower entropy regions are more intimately involved in complex problem solving processes compared to higher entropy

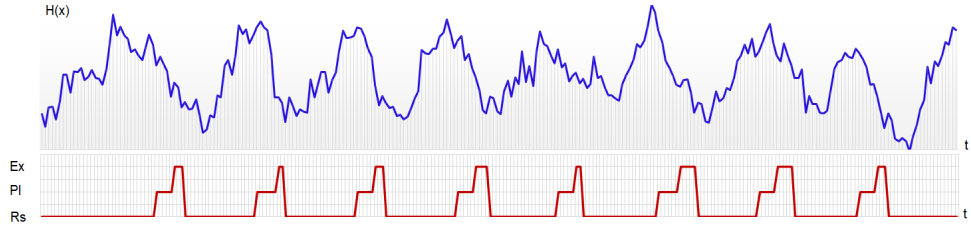


Figure 5.6: Expert Player: Dynamic Entropy fluctuations of a low entropy region, Right Lingual during session-3 of an expert player, Subject-175 (blue). *Pl* shows the time instances when the subject is doing Planning, *Ex* shows the execution and *Rs* shows, when the subject is back to the resting state, in the red plot.

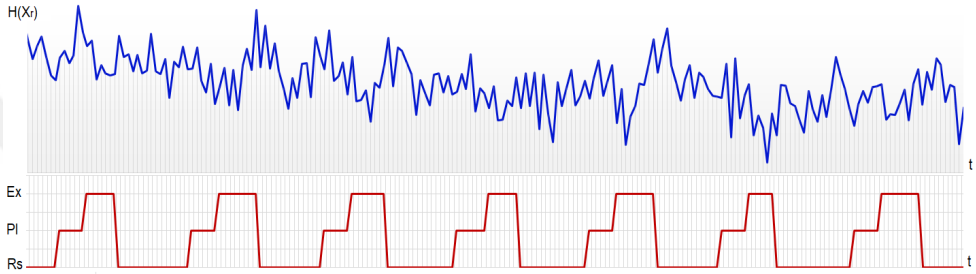


Figure 5.7: Novice player: Dynamic Entropy fluctuations for a low entropy region, Right Lingual during first session for a novice player, Subject-881 (blue). *Pl* shows the time instances when the subject is doing Planning, *Ex* shows the execution and *Rs* shows, when the subject is back to the resting state, in the red plot.

regions. Information theory provides a rigorous background for relating the information content and the regularity of a signal to the entropy measures.

First, Shannon entropy quantifies the *average amount of information* required to represent an event. If the state of a system or the outcome of an event is deterministic, or if it is very likely to occur, not much information is needed to describe it. Therefore, the entropy of such events are relatively low. Thus, measuring the entropy of a brain region give us the information about its state.

Second, Shannon entropy provides a quantitative specification of *disorder* of a dynamic system or *randomness* of an event. Therefore, increased system entropy is an indication of increased uncertainty. Thus, high entropy implies randomness or irregularity in the system. Conversely, a low entropy value indicates greater predictability of the system behavior or outcome.

In this study, we observe that the static and dynamic entropies are consistent with each other. Both entropies are relatively low in **superior parietal, occipital** regions, **precuneus, temporal** regions, **angular** gyrus, **prefrontal** regions, **lingual, inferior parietal, precentral** and **postcentral** areas, **cuneus, calcarine, paracentral, anterior cingulate** and **fusiform**, while the subject plays the TOL game.

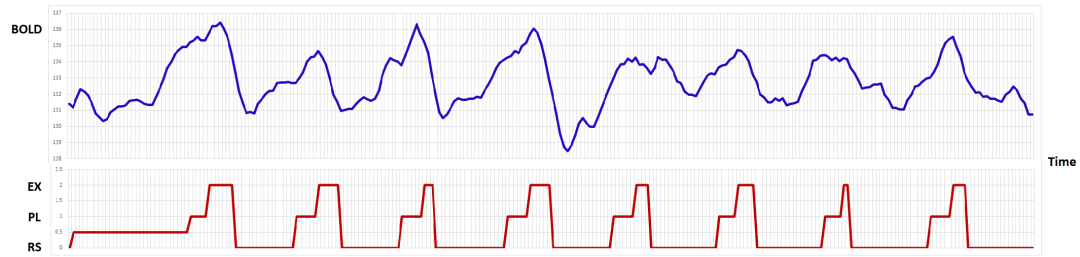


Figure 5.8: BOLD time series and CPS task correlation for Subj-146 sess4 (Expert player) for an activated region Right Superior Parietal during TOL problem-solving. *Pl* shows the time instances when the subject is doing Planning, *Ex* shows the execution and *Rs* shows, when the subject is back to the resting state, in the red plot

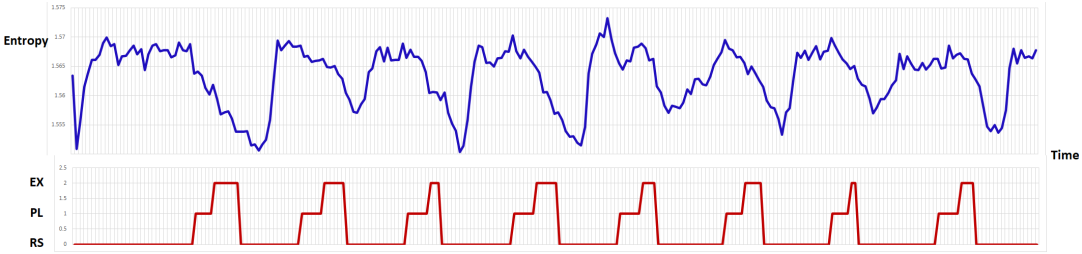


Figure 5.9: Entropy values and CPS task correlation for Subj-146 sess4 (Expert player) for an activated region Right Superior Parietal during TOL problem-solving. *Pl* shows the time instances when the subject is doing Planning, *Ex* shows the execution and *Rs* shows, when the subject is back to the resting state, in the red plot

The experimental studies in the literature, show that the active regions of complex problem solving task are visuospatial processing (**superior parietal** cortex) and executive processing such as working memory (**prefrontal** regions), planning (**prefrontal** cortex and **basal ganglia**), and error detection (**anterior cingulate**) [147]. It should be noted that some of the studies in the literature focused only on certain regions of the brain and ignored other parts.

The **prefrontal cortex** is suggested as an important part of the cortical network, especially involved in planning. Newman et al. proposed that left and right prefrontal cortices were equally involved during the solution of complex TOL problems. They suggested that the right prefrontal cortex might be connected to the plan generation, while the left part might be more related to the plan execution [150]. In another TOL study, considerable activation in the **middle frontal** gyrus and the adjacent part of the **inferior frontal** sulcus, the **precentral** cortex, and the **anterior cingulate** gyrus was observed [161].

The involvement of **superior parietal** areas are due to the spatial processing. The involvement of the **right superior parietal** region is due to the visuospatial attention. Note that, both regions necessary for planning. On the other hand, the **left parietal** re-

gion is linked to visuospatial working memory processing. Also, there are findings for the participation of the **superior posterior parietal** region in visuospatial processing, including the maintenance and transformation of internal representations [150].

Recent evidence suggests that the **precuneus** region is involved in various functions, including episodic memory retrieval, integration of information relating to perception of the environment, mental imagery strategies, visuospatial imagery [162–165]. The **occipital** lobe is the visual processing center of the brain.

Fincham and Anderson et al. have formalized planning behavior in the Tower of Hanoi (TOH) task as a computational model within the adaptive control of thought–rational (ACT-R) cognitive modeling framework (TOH can be seen as a similar problem-solving measure to TOL) [166]. They found that the regions, which differentially responsive to goal-processing operations distribute among **prefrontal cortex**, **parietal cortex** (**superior parietal, precuneus, cuneus, inferior parietal and angular gyrus**), **cingulate gyrus**, and subcortical structures.

Unterrainer et al. proposed that during TOL task, good problem solvers showed increased activation in **right superior temporal** region and **inferior parietal** region, which may reflect more visuospatial attentional processing, a more intensive recourse on stored visual information needed for the task. During the execution phase, incorrectly solved trials elicited more activations than correctly solved problems in **bilateral premotor** and **dorsolateral prefrontal cortex**, as well as in **left precuneus** and **temporal region** [162]. They supposed that **left parietal** region reflects changes to the problem representation, and **posterior parietal** region tracks changes in problem representation. The increase in **right inferior parietal** activation may reflect more visuospatial attentional processing. According to [147], the **anterior cingulate** has been linked to error detection. It associated with setting goals and error processing .

In a TOL study [162], it was seen the activation of **right parahippocampal** and **lingual** gyrus during execution. It was supposed that it reflects the increased demands on spatial working memory.

Comparison of the results of the above experimental studies with our entropy based approach indicates a high overlap between the low entropy regions and active anatomical regions observed by experimental neuroscience, for TOL problems. Therefore, we can infer that regions with low entropy correspond to anatomical regions that are active during complex problem solving.

5.2.4 Static Entropy Analysis for Planning and Execution Subtasks

One of our goals is to compare the activation of brain anatomical regions during the planning and execution subtasks. For this purpose, we calculate static entropy of each anatomical region on subject, session, and subtask basis. We estimate the static entropy of planning and execution in two sets of experiments.

In the first set of experiments, we use all the voxels to estimate the representative time series for each anatomical region during the planning and execution phases, separately, and, then, estimate the static entropy measures. As it is observed from Figure 5.10, static entropy measures for planning phase is relatively low compared to that of the execution phase.

However, it is well-known that there are some voxels and/or anatomical regions, which do not contribute to the underlying cognitive tasks, namely, planning and execution phases of TOL problem-solving. While contributing to some other brain processes, these irrelevant regions/voxels may hide the tasks under investigation and act as a background noise. There are many techniques to eliminate the redundant voxels/regions in fMRI literature. One popular technique is called ANOVA. This technique estimates and ranks the mutual information between a specific cognitive task and the voxel time series. Then, the most informative voxels, with the highest mutual information values are selected from the ranked list and the rest is eliminated. It is shown that when the most informative voxels are selected by ANOVA, the brain decoding performances for planning and execution phases are substantially improved [167]. Based upon this finding, it is assumed that dropping the voxels with small mutual information scores reduces the noise generated by the irrelevant voxels.

In the second set of experiments, we estimate the static entropy values based on the most informative voxels, selected by ANOVA. For this purpose, we estimate mutual information of each voxel time-series, when the subject is performing both planning and execution subtasks. Then, we rank the voxels according to their mutual information scores. For each time instance, we select 25.000 voxels out of 185.000. We calculate the representative voxel time series of each anatomical region by averaging only the most informative voxels, selected by ANOVA. Then, we estimate static entropy measures of the planing and the execution phases for all subjects (Figure 5.11). When we compare Figure 5.10 and Figure 5.11, we observe that the difference between the static entropy measures of planning and execution phases are more accentuated when we eliminate the irrelevant voxels by ANOVA. This observation reveals that employing only the most informative voxels decrease the static entropy of the planning phase compared to that of the execution phase. This fact can be attributed to the elimination of noise generated by the irrelevant voxels, during the selection of the most informative voxels according to the mutual information scores using ANOVA.

In the third set of experiment, we repeat the same method as described above for voxel selection and static entropy estimation for planning and execution phases. However, this time, we just employ the successful runs. As it is observed from Figure 5.12, the difference of static entropy measures between the planning and execution subtasks are even more accentuated in the successful runs. Planning subtask has much lower entropy measures compared to the execution subtask in all of the low entropy anatomical regions.

We also examine the static entropy of anatomical regions for unsuccessful runs. Figure 5.13 shows the anatomical regions that have lowest entropy values for planning and execution tasks for unsuccessful runs based on most informative voxels. As ob-

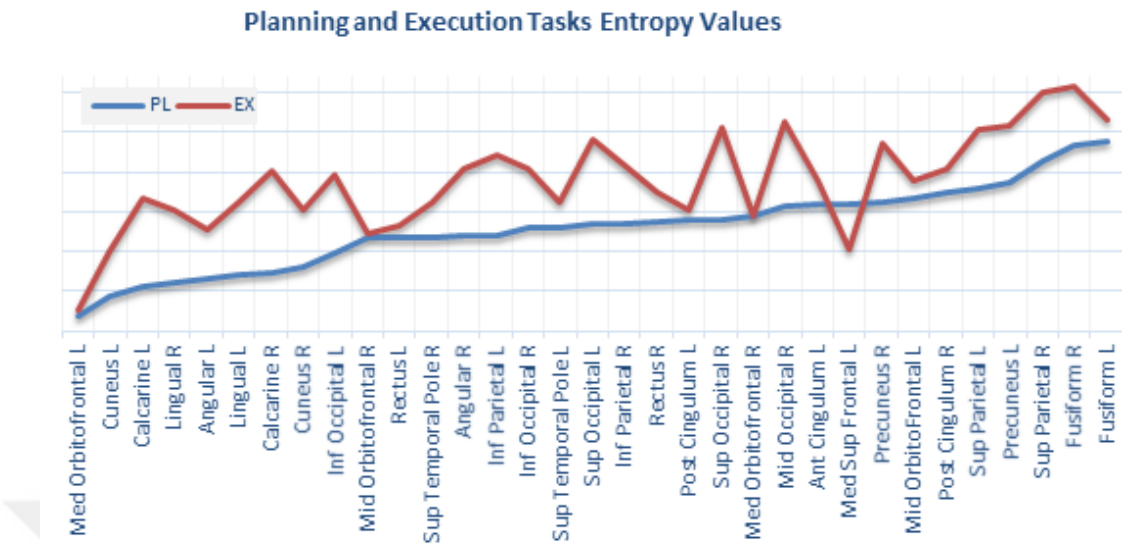


Figure 5.10: The lowest static entropy regions for planning and execution subtasks based on all voxels. The anatomical regions are sorted with respect to the lowest planning phase entropy values. Static Entropy estimation was made using the region representative time series averaged over all the voxels which resides in an anatomical region.

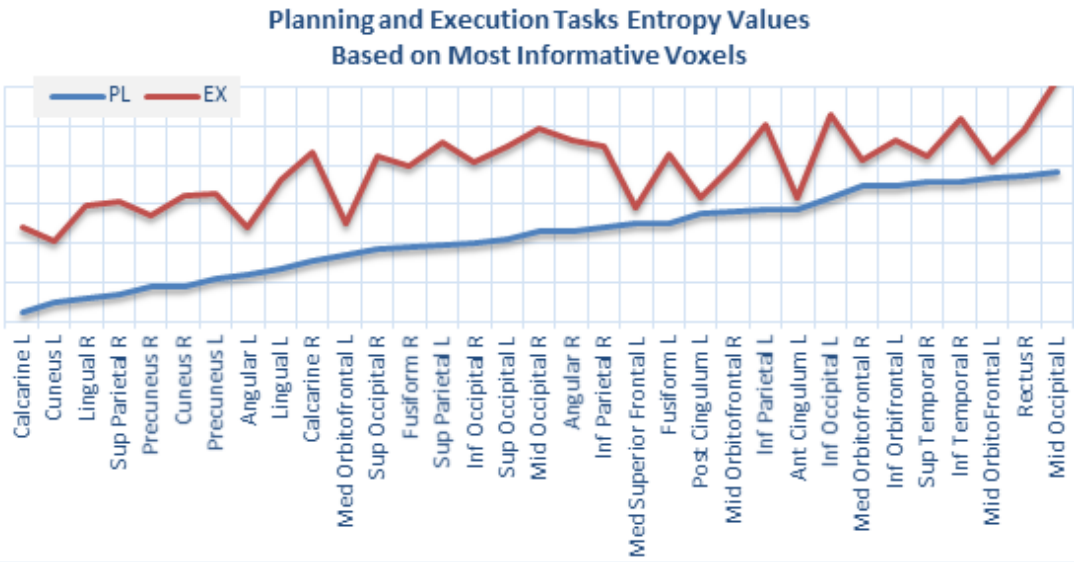


Figure 5.11: The lowest static entropy regions for planning and execution subtasks based on most informative voxels, selected by ANOVA method. The anatomical regions are sorted with respect to the lowest planning phase entropy values. Static Entropy estimation was made using the region representative time series by averaging only the most informative voxels, selected by ANOVA.

served from this figure, the difference of static entropy measures between the planning and execution subtasks is reduced in the unsuccessful runs.

The brain regions with the lowest static entropy measures for planning and execution subtasks based on most informative voxels are visualized in Figure 5.14 using Brain Net Viewer [168].

Low entropy regions in planning phase are **superior** and **inferior parietal** lobule, **precuneus** and **cuneus, calcarine, lingual, orbitofrontal** cortex, **angular** gyrus, **occipital** area, **left anterior** and **posterior cingulum, fusiform, superior** and **inferior temporal** pole. These results are mostly compatible with the results of various fMRI studies on TOL problem-solving [147, 161, 162, 169]. In execution phase, the regions that have low entropy values are **superior-inferior parietal, precuneus, cuneus, calcarine, angular** gyrus, **med-orbitofrontal** gyrus, **left posterior** and **left anterior cingulum, left medial superior frontal** gyrus, **lingual, superior temporal, right superior-inferior occipital** gyrus, and **right inferior parietal** cortex.

This analysis reveals that the static entropy of planning task is lower than that of the execution task. The difference between planning and execution entropy is higher for most informative voxels than all voxels. This result is compatible with the previous BOLD analysis study [167]. Since plan generation requires the participation of numerous cognitive processes, it would lead to more coherent neural processing than the execution task [170]. The difference between planning and execution entropy is higher for successful runs than that of the unsuccessful runs. This result can be interpreted as follows: The informative anatomical regions of an expert player is more organized compared to that of a novice player.

5.2.5 Static Entropy Analysis for Expert and Novice Subjects

In this section, we discuss the difference between expert and novice players in problem solving skills and examine regional entropy figures during TOL problem solving for expert and novice problem solvers.

5.2.5.1 Expert-Novice Differences in Problem-Solving Skills

It has been shown that expert and novice players think and solve problems in different ways. Although different abilities often outweigh in different types of problems (for example, players with higher spatial ability have higher accuracy on the TOL problems [147]), studying this distinctions can help explain why some people are more successful at problem-solving than others.

The ability to solve problems successfully depends on several factors. Studies on problem-solving have shown that experts differ in problem-solving skills mainly from novices in terms of the following attributes:

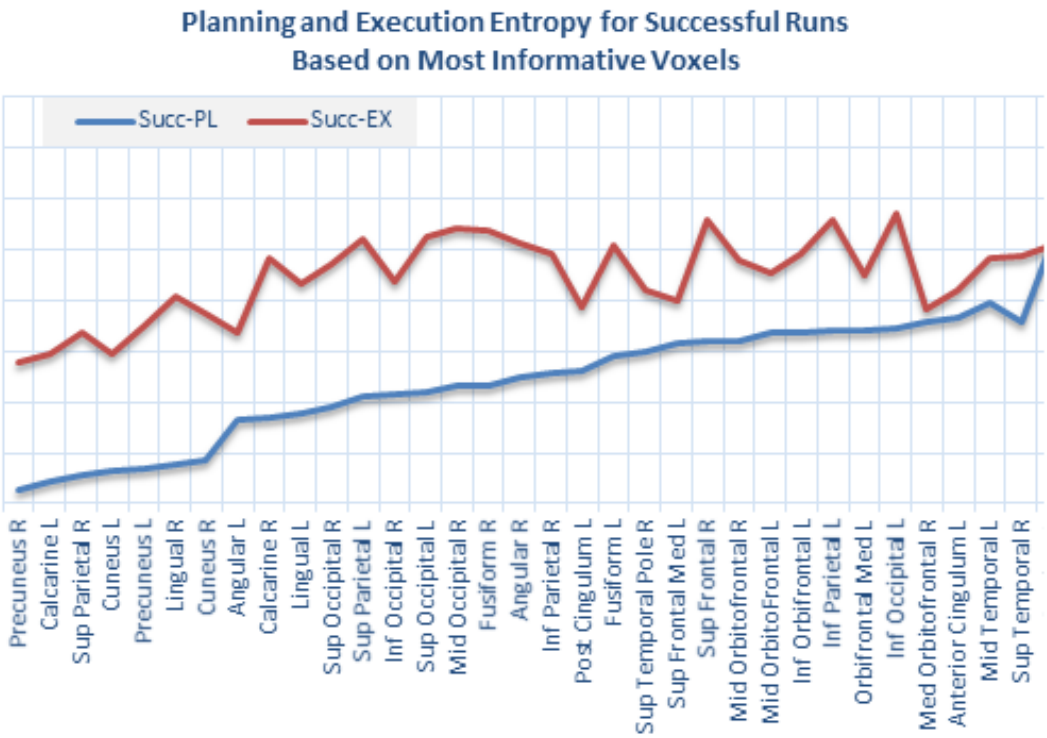


Figure 5.12: The lowest static entropy regions for planning and execution subtasks for successful runs based on most informative voxels, selected by ANOVA. The anatomical regions are sorted with respect to lowest planning phase static entropy values. Static Entropy estimation was made using the region representative time series,by averaging only the most informative voxels, selected by ANOVA.

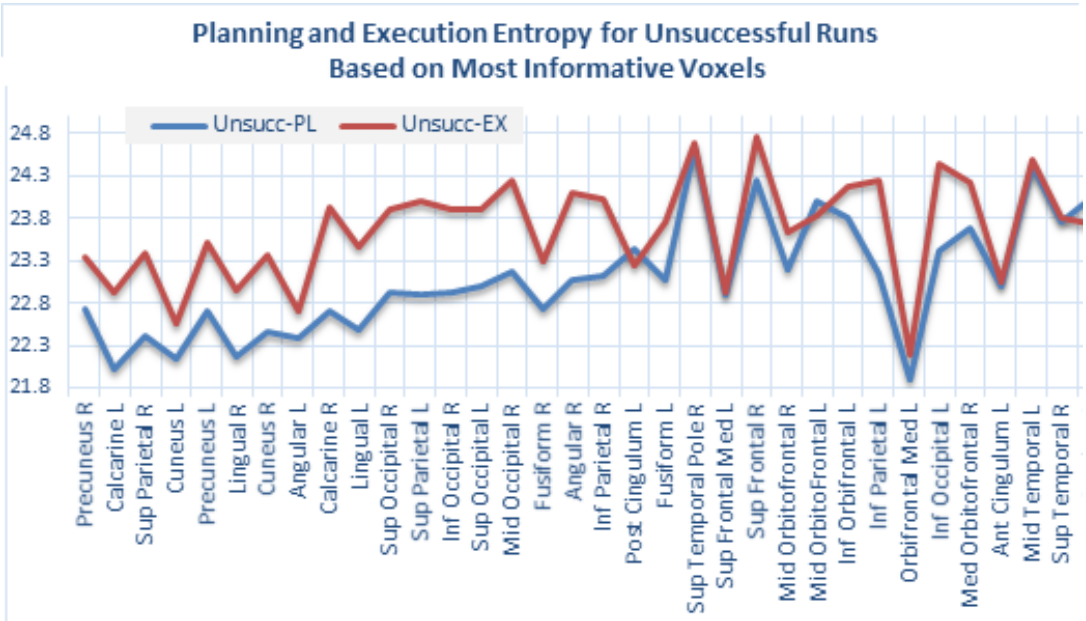


Figure 5.13: The lowest static entropy regions for planning and execution tasks for unsuccessful runs based on most informative voxels (sorted by minimum planning entropy for successful runs).

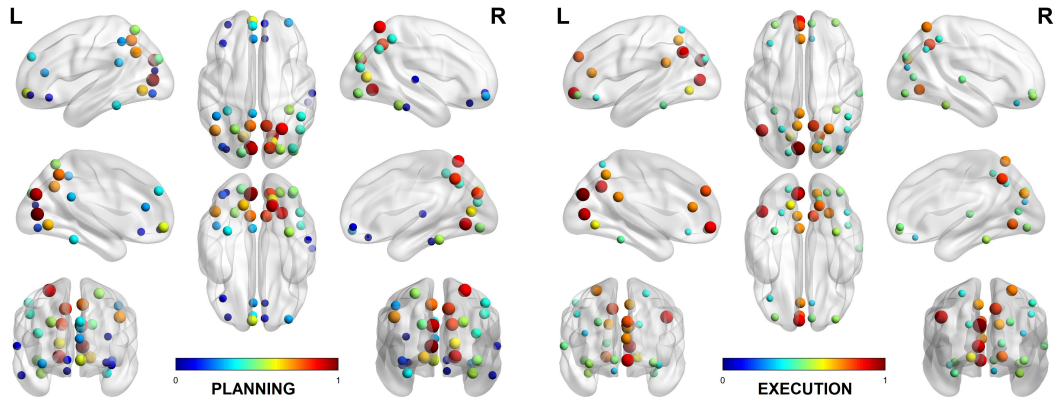


Figure 5.14: Visualization of anatomical regions with the lowest static entropy measures for planning and execution phases based on most informative voxels, selected by ANOVA. Hot colors represent relatively low static entropy regions. The size of the circles are inversely proportional to the static entropy measures

Domain specific knowledge: Expert players are supposed to have more domain-specific information to facilitate problem-solving compared to novice players. Also, they can learn new domain-relevant information more easily than novices. Additionally, they can distinguish large, meaningful patterns in their domain. Then, they apply domain-specific strategies better than novice players. They have more declarative knowledge (i.e. descriptions of the world including facts, theories, events, and objects) and procedural knowledge (i.e. how to perform particular actions to accomplish task goals). Experts' knowledge is also well-organized in ways that organize, represent, and interpret information. That means their knowledge turns into knowledge structures (i.e., chunks and schemas) that make it easy to categorize and construct a mental representation of the problem.

Memory representations: The quality of the memory representations is different between expert and novice problem solvers. Experts perceive information that is conceptualized more on the level of principles. The thought processes of experts reveal more complex and sophisticated representations of problems. In addition, experts is supposed to have better short and long term memory than novices, based on superior memory organization rather than volume.

Problem space: Problem-solving is characterized as a search through a problem space. The structure of the problem space determines the possible methods that can be used for problem-solving. Problem space includes various components that are used in the process of finding a solution to a problem. Expert problem solvers are supposed to explore a broader problem space than novice problem solvers.

Mental schemas: Experts possess well-elaborated schemas (i.e. organized representations of things or events) that provided a framework for meaningful interpretation of information, also, enables pattern matching and categorization of problems [147].

These schemas sometimes can lead directly to a path to the solution. Saariluoma [171] claimed that the process of selecting a move by an expert can be defined as a series of pattern recognition- restructuration cycles that make it possible to find solutions in the problem area with only a limited search.

Chunking information: It is suggested that experts chunk information and can access related chunks of knowledge from long-term memory. Thus making their problem-solving procedure more efficient. The chunking theory (Chase & Simon, 1973 [172]) proposed that at the core of expertise lies the ability to rapidly recognize important problem features. These features, internally stored as chunks, act as access points to semantic long-term memory. According to the Template Theory (Gobet & Simon, 1996 [173]), templates are "chunks" that evolve into more complex data structures. They can expedite search by allowing information to be stored into the long-term memory quickly; letting a search in the template space in addition to a search in the move space. Chunks are linked to other information stored in the long-term memory, such as moves, plans, tactical motives [174]. Studies describing the cognitive processes of experts and novices during the chess game proposed that experts do not consider plays farther ahead than the novice; rather, experts choose among vastly complex moves. The expert can chunk relevant information, while novices envision single pieces of information.

Solution Strategy: The implementation of general problem-solving strategies, also called *heuristics*, is the key to problem-solving expertise [175]. Heuristics are cognitive rule of thumb, derived from experience, which can be used as guides in problem-solving processes. It is suggested that the number of heuristics is constant across skill levels. In the literature, two main strategies for inference control are proposed: forward reasoning and backward reasoning. In forward reasoning, one begins with the givens, then, works directly toward the goal. In backward reasoning, one starts at the goal, then, works toward the beginning. However, there is no full consensus in the literature on what strategy experts or novices are using. Some of the researchers propose incompatible theories to explain the search strategies used for solving problems. A more commonly reported finding (Larkin et al. (1980) [176], Chi (1981) [177]) is that novice problem solvers in physics use backward inferences as a search technique, while experts use forward inference [178]. In fact, most studies on problem-solving suppose that experts are more likely to use a *working-forward strategy*, whereas novices are more likely to use a *working-backward strategy*. Template and Chunk theories also support the idea that experts perform the search working forward by recursively applying the pattern recognition processes in the internal representation [174]. The use of this strategy mostly depends on having a deep understanding of a problem to be able to construct a correct concrete representation from the problem statement [179]. This is made possible by the fact that experts can classify problems in terms of basic principles. They have knowledge about approaches that work for specific domains.

Metacognition: Metacognition is the self-awareness, and the ability to monitor and control one's mental processing. It is supposed that experts are more likely to be able to plan their solutions at a descriptive meta-level than novices. They can organize concepts in terms of deeper abstract principles [177], whereas novices organize

Table 5.1: Comparison of durations during planning and execution phases for successful and unsuccessful sessions.

TOL Experiment	Number of Time instances		
	Planning Phase	***Execution Phase***	***Pl-Ex Difference***
Successful Sessions	3.857	3.287	570
Unsuccessful Sessions	3.800	4.005	-205

concepts in terms of surface features of the problem. Experts also are assumed that to have strong self-monitoring skills. They can efficiently monitor their problem-solving progress by refining and correcting solutions. They are more aware of when they make errors, why they fail to comprehend and when they need to check their solutions [180].

Time allocation: Experts spend more time quantitatively analyzing and evaluating a problem before starting to execute the problem [180]]. They tend to devote most of their problem-solving time to the planning phase to figure out how to represent a problem and find the optimum solution path. Novices, on the other hand, tend to spend relatively more time in the execution phase. It has been shown that, longer planning times increase the likelihood of optimal solutions. Thus, longer planning times result in superior performance [147].

During the TOL game, if the planning phase is successful, that means a solution plan is successfully developed, the problem solver directly implements the solution plan that is stored in working memory. In this case, the execution task is accomplished in a relatively short time. However, the planning task may not always be complete or accurate. For complex problems that require the implementation of multiple sub-goals, the solution process might be an intermix of the planning and execution tasks. As a result, re-planning is required during execution to make corrections to the preliminary plan. Thus, we expect longer times for the completion of execution-replanning-execution processes.

In our TOL experiment, we observe that for most of the puzzles, the planning duration is longer than the execution duration during the successful runs. This observation implies that the expert players spent more time planning the solution than implementing the solution plan. Conversely, novice players spent less time on planning and more time on the plan execution. The comparison of the number of time instances recorded during planning and execution phases for successful and unsuccessful runs is shown in Table 5.1.

5.2.5.2 Active Brain Regions for Experts and Novices During TOL Problem-Solving

Previous studies related to TOL problem-solving have reported extensive prefrontal, parietal, cingular, and striatal brain activity; especially activations in the dorsolateral prefrontal cortex, anterior cingulate, posterior cingulate, and bilateral posterior and inferior parietal cortices [161, 169, 181, 182]. The majority of the studies on problem-solving are focused on the planning phase of problem-solving. Lazeron et al. [161] reported the following areas of fMRI activation during the TOL task planning phase: Middle frontal gyrus, Inferior frontal gyrus, anterior and posterior Cingulate gyrus, Cuneus and Precuneus, Supramarginal gyrus left, Angular gyrus left, insula left, and bilateral Occipital gyrus.

Cazalis et al. [169] show that healthy subjects with different levels of performance in the TOL planning task exhibit different patterns of brain activation. While all subjects show significant activation in the dorsolateral prefrontal cortex (DLPFC), the anterior and posterior cingulate areas, and the parietal cortex, good players showed more activation in the left DLPFC than standard players. The standard players tended to show increased activation of the anterior cingulate region.

Unterrainer et al. [162] reported that during the TOL task, experts showed increased activation in inferior parietal and right superior temporal regions, which may reflect more visuospatial attentional processing. During the execution phase, incorrectly solved trials showed more activations than correctly solved ones in the left inferior parietal, bilateral premotor, and bilateral prefrontal cortex, as well as in the anterior cingulate cortex bilaterally, left precuneus, and parahippocampal gyri. They supposed that the left parietal region reflects changes to the problem representation, and the posterior parietal region tracks changes in problem representation. The increase in right inferior parietal activation may reflect more visuospatial attentional processing. Experts also showed increased activation in inferior parietal and right superior temporal regions. For correctly solved trials, the performance level was significantly positively correlated with activation in the right posterior thalamus. During the planning phase, significant bilateral activations can be seen in the prefrontal cortex, in parietal regions, in the anterior and posterior cingulate cortex, the basal ganglia, and the thalamus. In the planning condition, experts showed increased activation in the right DLPFC as well as in the right superior temporal and inferior parietal regions.

In our study, we estimated entropy values of brain regions during the TOL task for expert and novice players separately. We observed that superior parietal and precuneus have lower entropy values (i.e. exhibit high activations) for both experts and novices. Also, we do not find any significant differences for correctly and incorrectly solved problems in the following low entropy regions: postcentral, middle and inferior occipital right, angular right, inferior and middle temporal right, precentral right, and cuneus left.

We investigated the anatomical regions that show different activation patterns for expert and novice players when solving TOL problems. We observe that the following regions have lower entropy values for experts than novices: left middle and superior

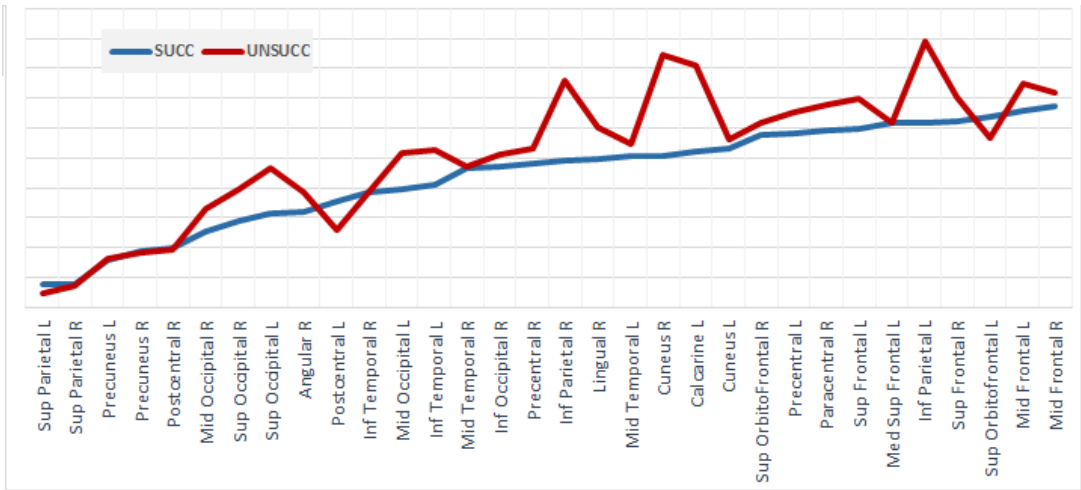
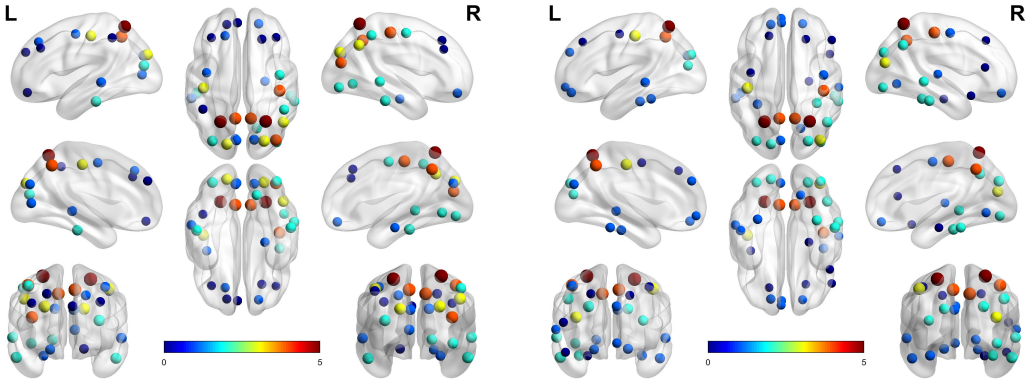


Figure 5.15: The anatomical regions with the lowest static entropy values of expert and novice subjects based on most informative voxels, selected by ANOVA method. The anatomical regions are sorted with respect to the lowest planning phase entropy values.

frontal, right inferior orbitofrontal, inferior parietal, posterior and middle cingulate, supplementary motor area, right olfactory, left postcentral, left angular, left supramarginal, superior and middle occipital, lingual, cuneus, left calcarine, left inferior temporal, left middle temporal pole, right hippocampus, hesch, amygdala, thalamus, left rolandic operculum, left pallidum, and right putamen. Novices have lower entropy values than experts in the following anatomical regions: left inferior- middle-medial orbitofrontal, anterior cingulate, superior temporal pole, fusiform, parahippocampal, left hippocampus, insula, caudate, right rolandic operculum, right pallidum, left putamen.

As the above list shows, the right or left parts of several regions differ for expert and novice players. For example, the right inferior orbitofrontal gayrus and the right hippocampus shows more activation for expert players, whereas, the left parts are more active for novices. For Rolandic opreculum and Pallidum, left parts are more activated for expert players, yet right parts show more activation for novices.

Figure 5.15 provides a comparison of the low entropy regions between strong and weak problem solvers based on most informative voxels, selected by ANOVA. The brain regions in this figure are visualized in Figure 5.16 using Brain Net Viewer [168]. Figure 5.17 and Figure 5.18 show the comparison of anatomical regions of experts and novices for planning and execution phases of the TOL game respectively.



(a) Experts (b) Novices

Figure 5.16: Visualization of anatomical regions with the lowest static entropy measures for expert and novice players based on most informative voxels, selected by ANOVA. Hot colors represent relatively low static entropy regions. The size of the circles are inversely proportional to the static entropy measures.

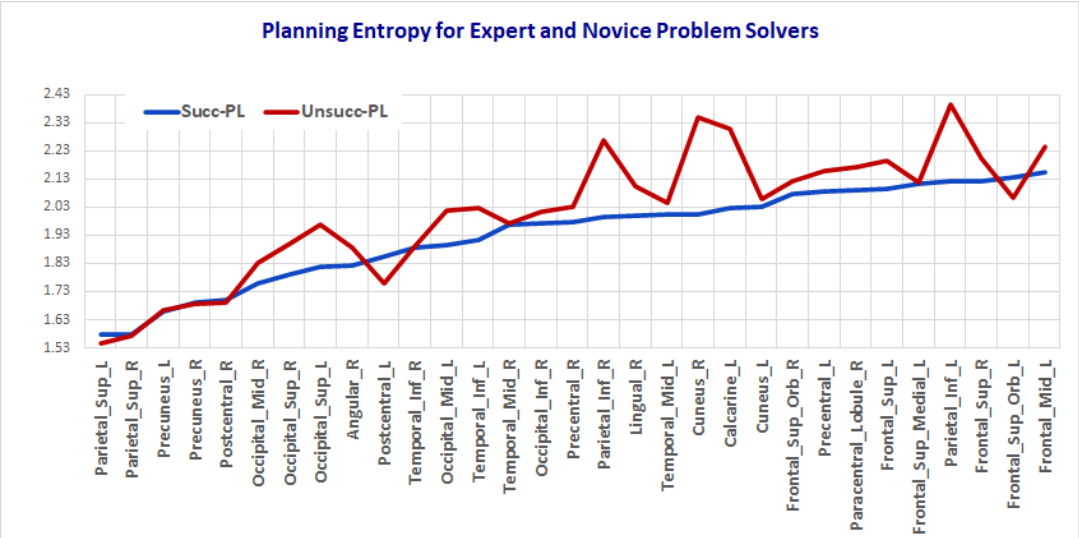


Figure 5.17: Low entropy regions for expert and novice subjects for **planning phase** of TOL game.

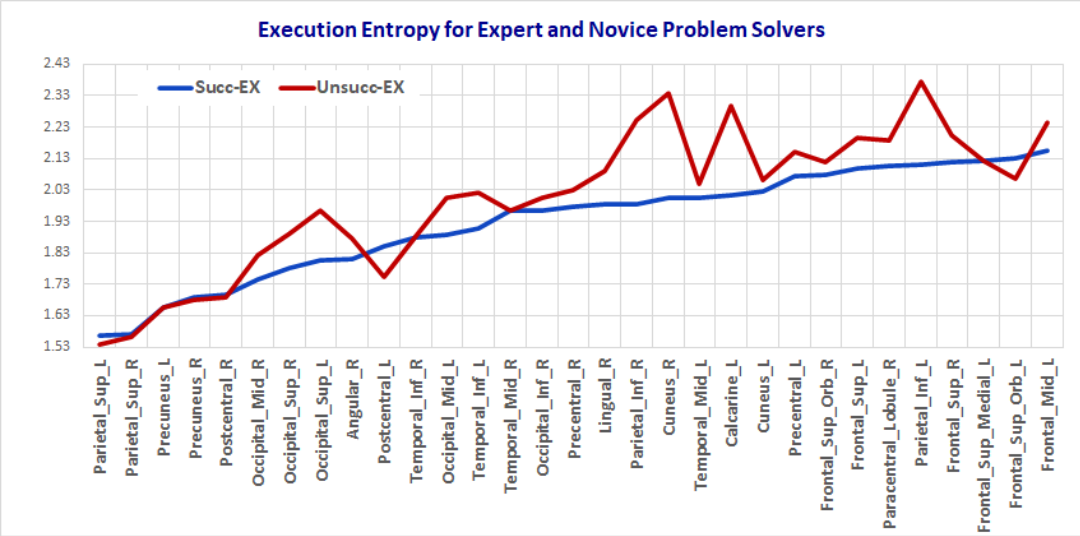


Figure 5.18: Low entrpy regions for expert and novice subjects for **execution phase** of TOL game.

5.3 Static and Dynamic Brain Networks Analysis

The methods introduced in Section 4.3, "Estimating Static and Dynamic Brain Networks", explain how we generate static and dynamic arc-weight matrices between brain anatomical regions, where Kullback-Leibler (K-L) measures represent the degree of co-activation between the pairs of the 90 anatomical regions. Therefore, each matrix has $90x90 = 8100$ connections for each static brain network. On the other hand, we estimate $90x90 = 8100$ connections for each time instant to represent dynamic brain networks. In the case of TOL experiment, the dynamic K-L matrices are generated for each subject, session and time-instances. Therefore, the number of K-L matrices we generate is: 18 subjects x 72 sessions x 590 time-instances = 764.640 and the number of K-L connections is 764.640 x 8100 = 6.193.584.000.

In order to show the "most significant" connections among the anatomical regions, the arc weights of K-L matrices are reduced by applying a threshold value. In order to display a visually meaningful matrice, we eliminate the K-L values, greater than 0.023, which is chosen empirically. The combined pair of connections are processed according to the lowest K-L distances and the most frequent connections occurred across all planning and execution time instances, to find out the most significant connections. Finally, the static K-L matrices are generated for planning and execution phases, separately.

Using the static K-L divergence estimation method, we observe that, for most cases, the number of minimum K-L distances between anatomical regions for planning phase, is higher than the number of minimum K-L distances for execution phase. Figure 5.19 shows the lowest K-L distances for planning phase across execution phase for a subject. This result is consistent with the experimental neuroscience,

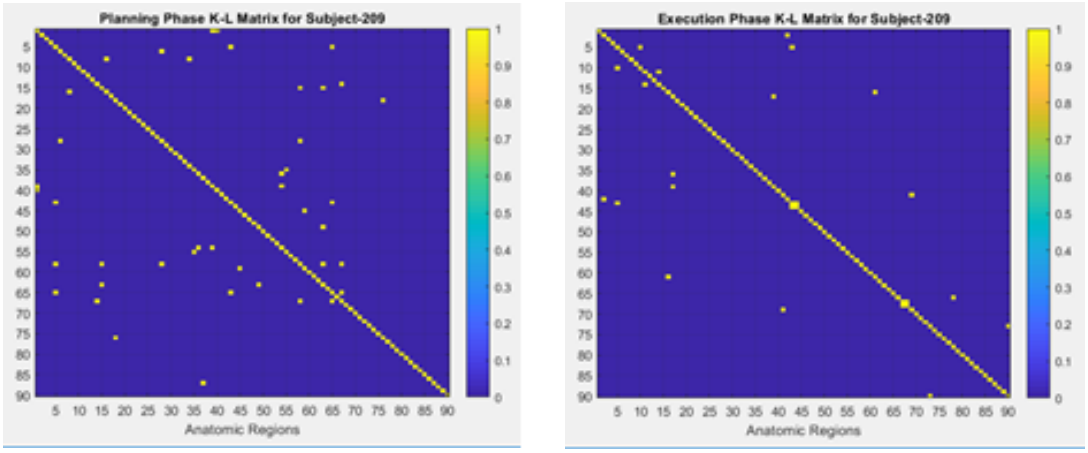


Figure 5.19: Static K-L divergences lower than a specific threshold (0.025) value between anatomical regions for *planning* phase (Left) versus the *execution* phase (Right) for subject-209.

which observes that planning is a more complicated process than the execution of the game [147, 159]].

We also estimated Pearson correlation matrices for each subject, session and time-instance, and compare them to K-L matrices. Since each region contains different number of voxels, Pearson correlation matrices are calculated using voxel distributions, instead of voxel intensity values. Figure 5.20 shows sample images of K-L distance matrix and Pearson correlation matrix values which below a specific threshold value (K-L distance $<= 0.023$, Pearson p-value $<= 0.01$ and abs(r) > 0.98. Pearson matrix is based on absolute r values. K-L distance values are normalized to 0-1 interval, then subtracted from 1 in order to emphasize the shortest distances, and also to provide the compatibility with Pearson correlation values. Since K-L distance and Pearson correlation values are different metrics, the specified threshold values could not be the same. As a result, the number of strong connections is slightly different. However, as can be seen in Figure 5.20, there are some similarities between the two matrices.

We also generated the K-L matrices for successful and unsuccessful sessions. For the classification of successful and unsuccessful sessions, we made an analysis based on the number of puzzles completed successfully in a session and how many movements made to complete according to the shortest path. The top 42 sessions that contains 756 puzzles are considered as successful and the rest are considered as unsuccessful.

It is well-known that K-L divergence is not symmetric. Therefore, the brain networks generated by K-L matrices are directed. In order to avoid the complexity in visualization introduced by the directed networks, we also generated the undirected version of K-L divergences, which takes the average,

$$D_{KL} = 0.5\,(D_{\mathrm{KL}}(P_k \parallel P_l) + D_{\mathrm{KL}}(P_l \parallel P_k)), \tag{5.1}$$

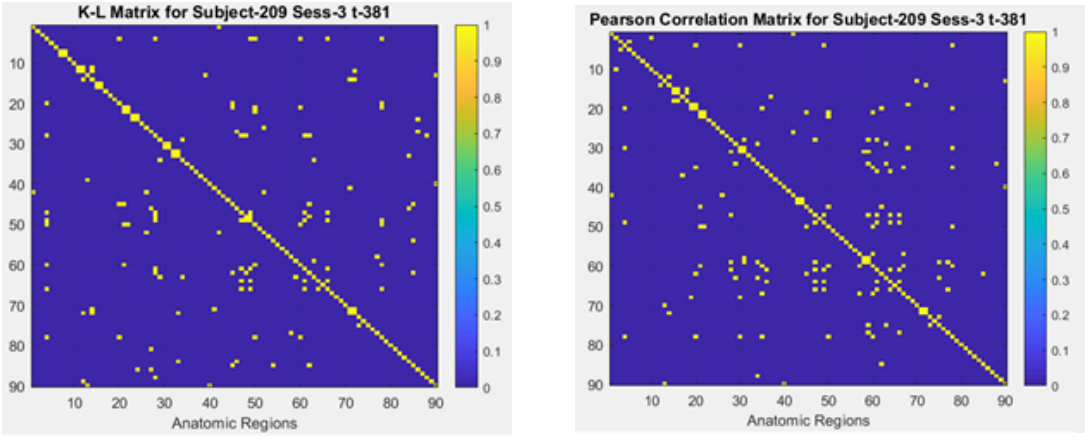


Figure 5.20: Static K-L divergences (Left) and Pearson Correlation matrix (Right) during a specific time instance for Subject-209 session-3.

as the arc weights.

5.3.1 Prominent Connections for Successful Sessions during the Planning and Execution Phases

Studies show that two main phases of problem-solving, namely planning and execution phases, involves different types of operations. As we mentioned in the previous sections, planning includes a set of operations for the construction of a problem representation and searching for the appropriate operators to solve the problem; and interacts with various cognitive processes such as abstraction, inference, analysis, synthesis, decision making. Execution phase involves the implementation of the planned solution, and it requires different neural activities compared to the planning phase.

In order to analyze the degree of connectivity among the anatomical regions during the the planning and execution phases, we select the most significant directed and undirected K-L connection matrices. Figure 5.21 and Figure 5.22 show prominent connections representing activity among anatomical regions for planning and execution phases for successful TOL problem-solving sessions, consecutively. Figure 5.23 and Figure 5.24 show the corresponding visualization of anatomical regions which have the smallest K-L distance measures. As can be seen in the figures, the execution phase connections depicts different patterns than planning phase network.

5.3.2 Planning and Execution Phases Prominent Connections for Unsuccessful Sessions

In this section, we compare the static brain networks for unsuccessful sessions using undirected K-L distance matrices. Figure 5.25 and Figure 5.26 show planning and execution phases undirected functional connectivity graphs for unsuccessful sessions,

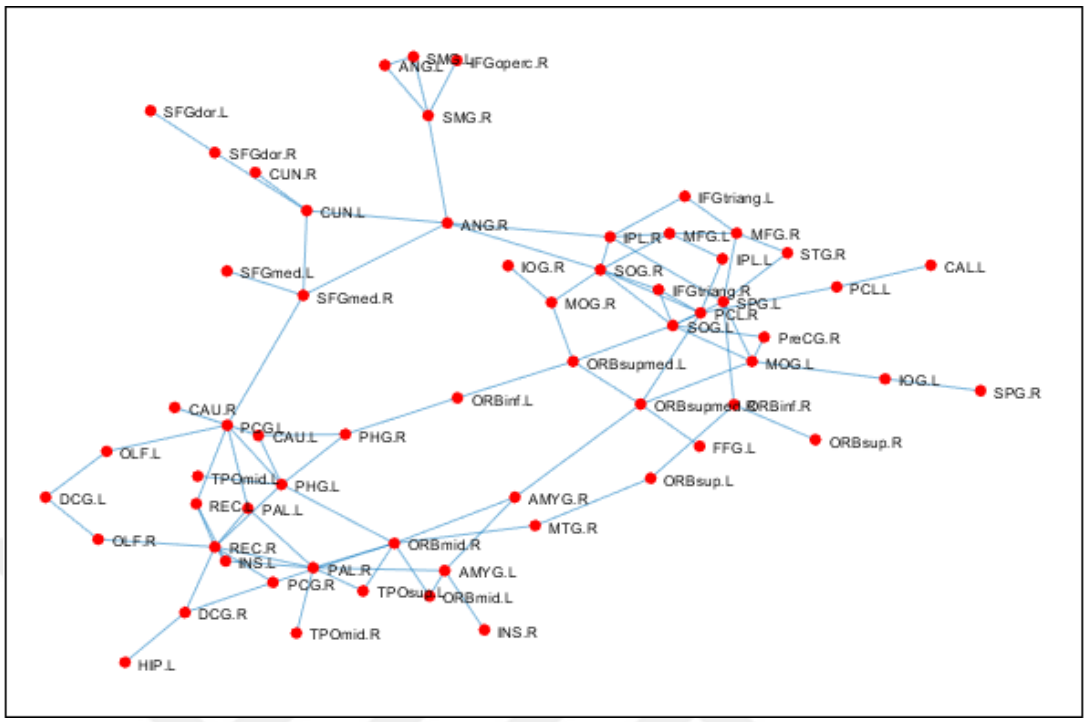


Figure 5.21: Undirected connection graph which shows most prominent connections for **planning** phase for **successful** TOL problem-solving sessions.

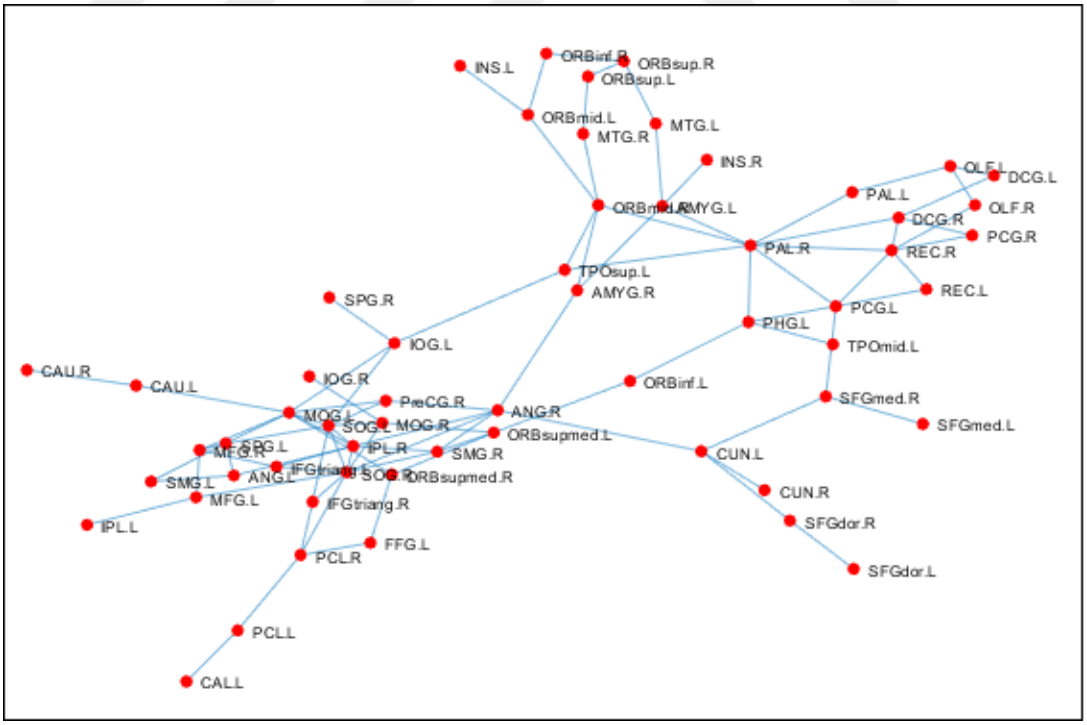


Figure 5.22: Undirected connection graph which shows most prominent connections for **execution** phase for **successful** TOL problem-solving sessions.

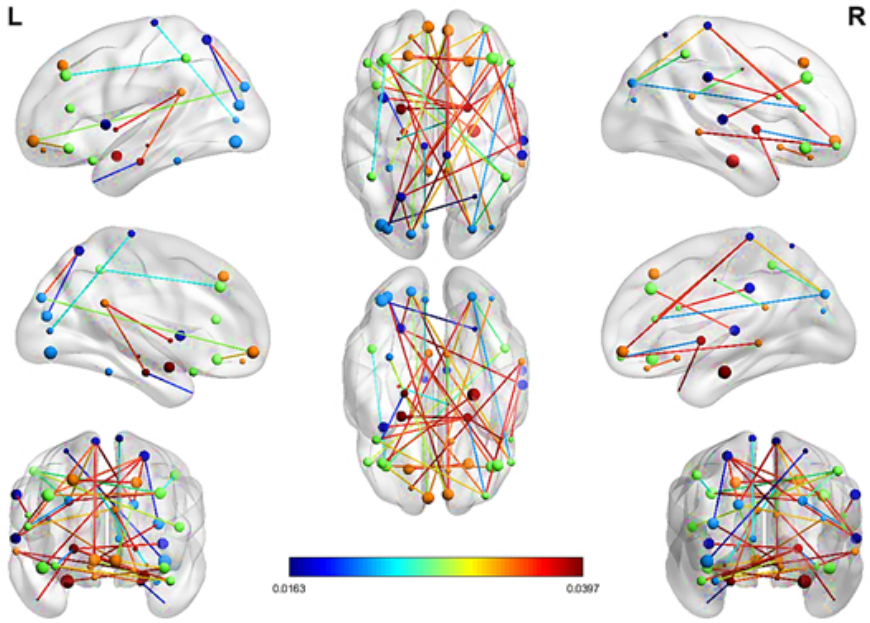


Figure 5.23: Visualization of anatomical regions which have shortest K-L distance measures for **planning** phase for **successful** sessions.

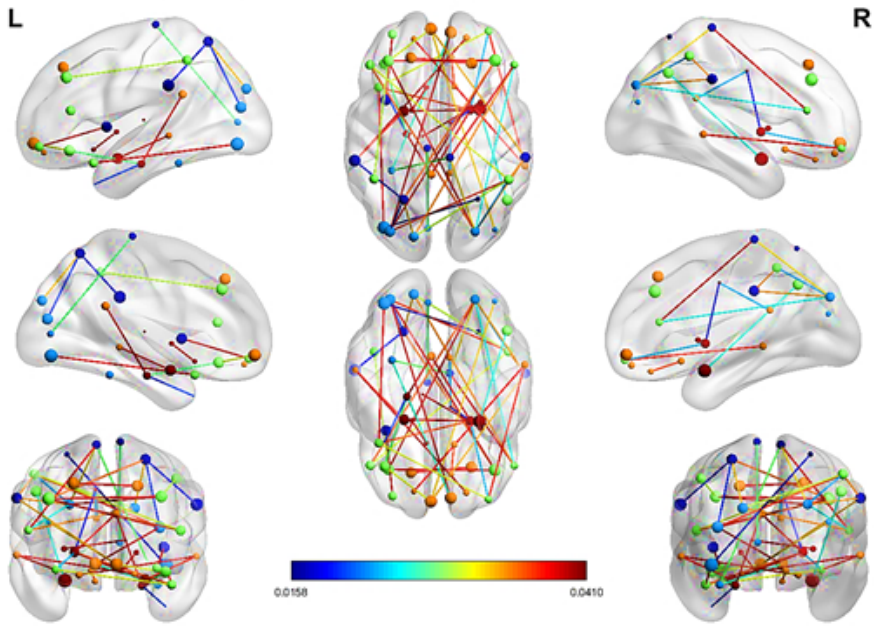


Figure 5.24: Visualization of anatomical regions which have shortest K-L distance measures for **execution** phase for **successful** sessions.

consecutively. Figure 5.27 and Figure 5.28 show the corresponding visualization of anatomical regions which have shortest distance measures for planning and execution phases for unsuccessful sessions, consecutively.

5.3.3 Analysis of Brain Network Properties

In this section, the co-activities among the anatomical regions are analyzed with the network properties, namely node degree, closeness centrality and betweenness centrality.

5.3.3.1 Node Degree

Node degree is the number of connections of a node, to the other nodes of the brain network. It quantifies the hub brain regions interacting with a large number of brain regions, thus, a node with high degree indicates its central role in the network [148]. For TOL experiment, the high degree nodes are mostly in the *frontal regions, parahippocampal gyrus, amygdala, occipital regions, caudate, insula, supramarginal gyrus*.

The high degree nodes for planning and execution tasks are:

- Planning: right parahippocampal gyrus, right amygdala, left inferior occipital gyrus, right insula, right middle frontal, inferior and superior medial orbitofrontal, right superior frontal medial and dorsal, left supramarginal, right caudate, left middle occipital gyrus.
- Execution: right amygdala, right parahippocampal, left inferior occipital, inferior and superior medial orbitofrontal, middle frontal, superior frontal medial, caudate, insula, supramarginal gyrus, left middle occipital, inferior parietal, angular gyrus.

Figure 5.29 shows that bilateral superior frontal gyrus, left superior frontal gyrus, left middle frontral gyrus, right orbitofrontal cortex, right supramarginal gyrus, right insula, left cuneus, left superior temporal gyrus and left inferior parietal lobule have more connections during planning phase than the execution phase.

5.3.3.2 Closeness centrality

Centrality measures node importance in a network. Closeness centrality estimates how fast information flows from a given node to other nodes. It indicates how close a node is to all other nodes in the network. It is calculated as the inverse of the sum of the lengths of the shortest paths between the node and all other nodes. For TOL experiment, the more central regions are in the *frontal, temporal and occipital lobes* respectively.

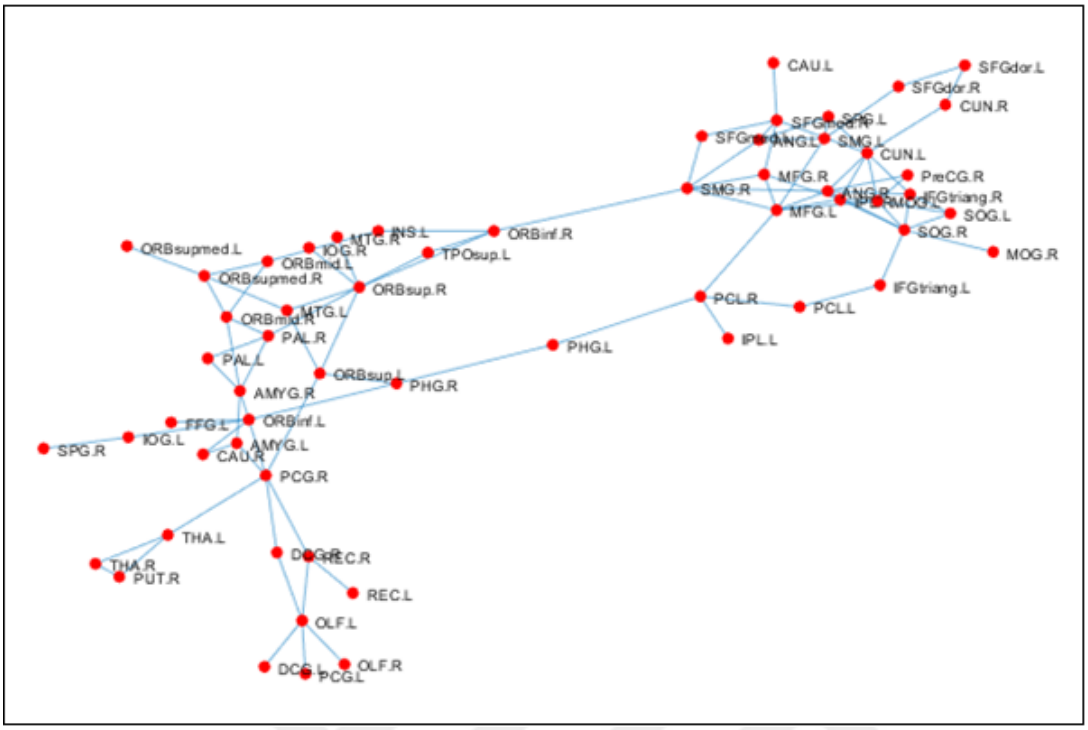


Figure 5.25: Undirected connection graph that shows most prominent connections for **planning** phase for **unsuccessful** TOL problem-solving sessions.

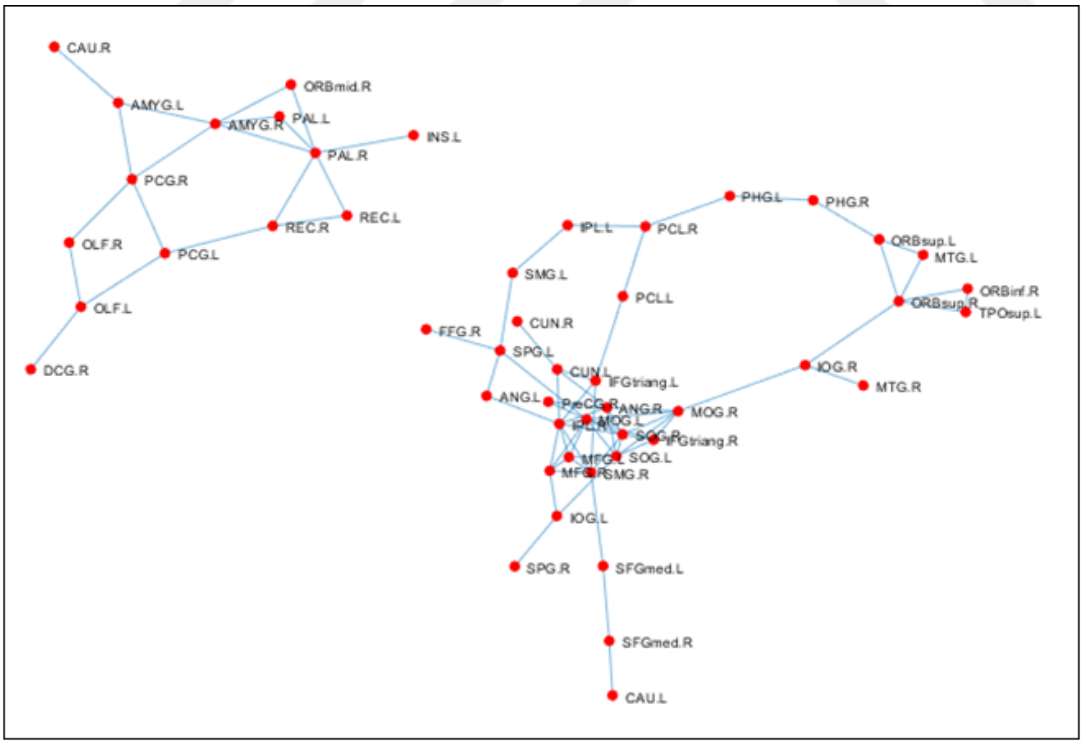


Figure 5.26: Undirected connection graph that shows most prominent connections for **execution** phase for **unsuccessful** TOL problem-solving sessions.

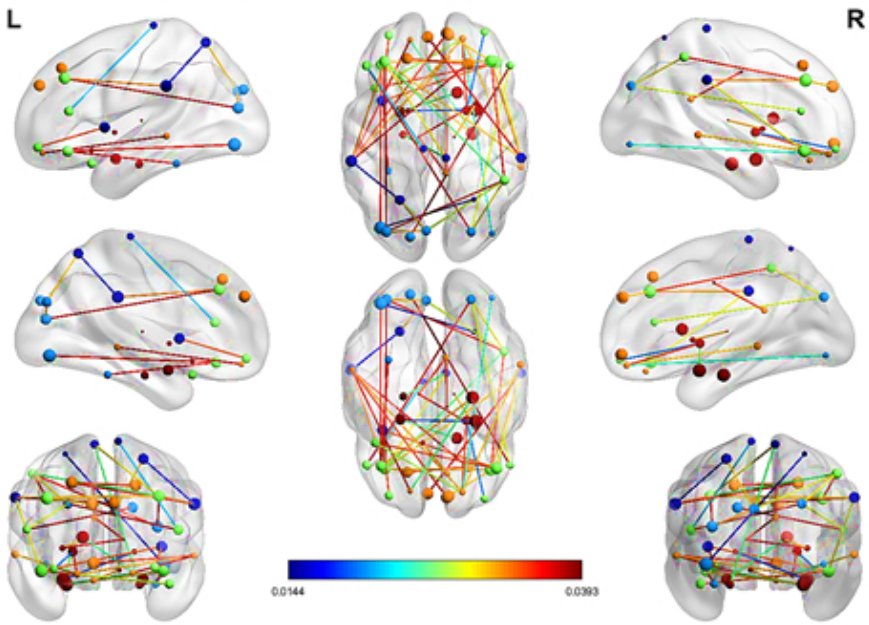


Figure 5.27: Visualization of anatomical regions which have shortest K-L distance measures for **planning** phase for **unsuccessful** sessions.

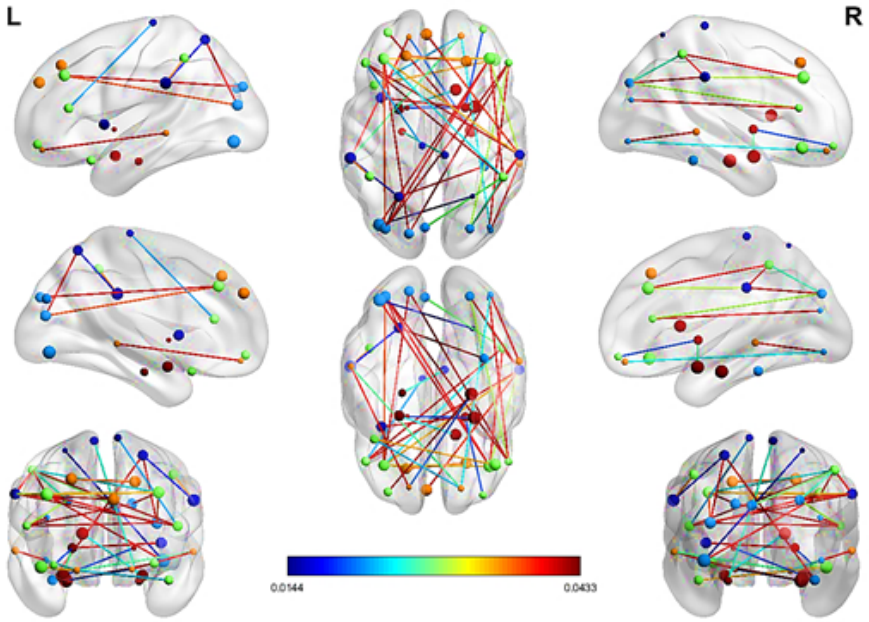


Figure 5.28: Visualization of anatomical regions which have shortest K-L distance measures for **execution** phase for **unsuccessful** sessions.

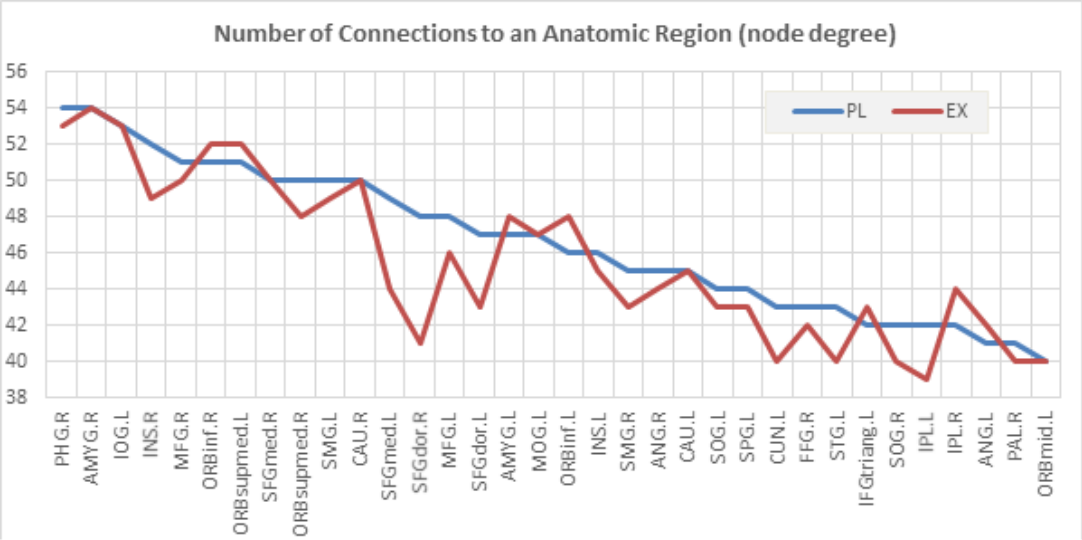


Figure 5.29: Number of connections with other regions to a region during planning and execution tasks.

- Planning: superior frontal gyrus medial right, amygdala, superior medial and inferior orbitofrontal, middle occipital left, superior parietal left, parahippocampal right, middle frontal, angular, superior occipital left, posterior cingulate gyrus.
- Execution: amygdala, superior frontal gyrus medial right, middle occipital left, superior medial and inferior orbitofrontal, superior parietal left, angular, caudate, parahippocampal right, superior occipital, middle frontal.

Figure 5.30 shows that, right middle frontral gyrus, left angular gyrus, left precentral lobule, right insula and right inferior parietal lobule are more central among the other regions during planning than the execution.

5.3.3.3 Betweenness centrality

Betweenness centrality is the fraction of all shortest paths passing through a given node in the network. Nodes with high betweenness centrality value are assumed to participate many of the shortest paths of the networks; thus, it plays an important role in the information flow of the network [148]. For TOL experiment, the regions that participated the shortest paths are *superior frontal, superior and inferior orbitofrontal, left posterior and anterior cingulate cortex, right amygdala, left superior parietal, right pallidum, right insula and left parahippocampal gyrus.*

- Planning: superior frontal gyrus medial right, posterior cingulate gyrus left, inferior-superior-medial orbitofrontal regions, anterior cingulate left, amygdala

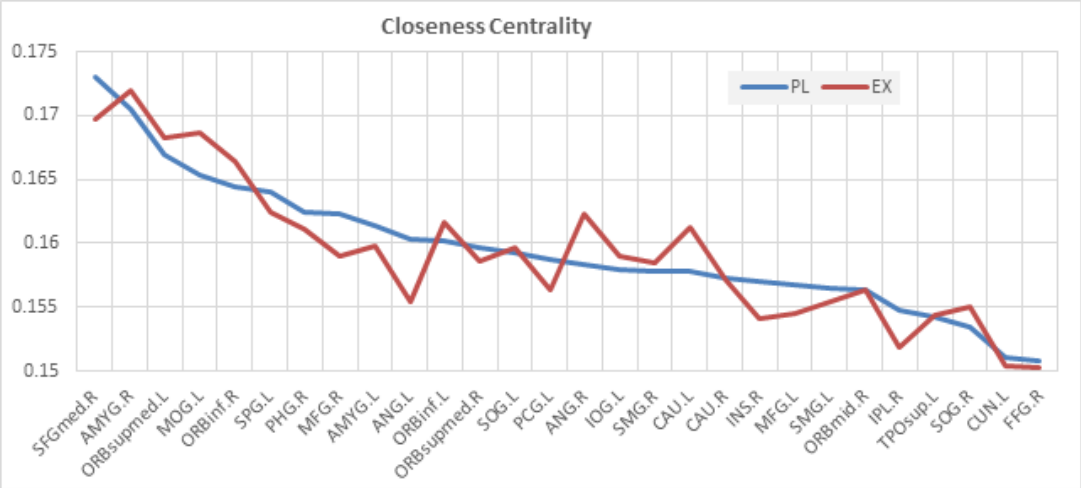


Figure 5.30: The regions that have high closeness centrality for planning and execution tasks.

right, superior parietal left, pallidum right, insula right, parahippocampal gyrus left.

- Execution: superior frontal gyrus medial right, inferior-superior-medial orbitofrontal regions, posterior cingulate left, anterior cingulate left, amygdala right, pallidum right, parahippocampal gyrus left, insula, middle occipital left.

Figure 5.31 shows that left superior parietal, left superior frontal (medial) and left putamen participate the shortest paths during planning than execution phase.

5.3.3.4 Common Prominent Regions According to Brain Network Properties

The common prominent regions in accordance with analyzed network properties and rough general functions of those regions are: orbitofrontal cortex (decision making), middle and superior medial frontal gyrus (cognitive control, attention, working memory), amygdala (emotion), parahippocampal gyrus (memory), caudate (procedural execution), insula (perception, emotion, cognition), supramarginal gyrus (language, perception, cognition), occipital gyrus (vision), anterior and posterior cingulate (error detection, emotional regulation, memory retrieval, attention), angular gyrus (number processing, spatial cognition, memory retrieval, attention), inferior and superior parietal (visuo-spatial processing, internal representations), left cuneus (visual processing), putamen (learning, movement), pallidum (reward, movement). These results are mostly compatible with the results of various fMRI studies on TOL problem-solving [147, 150, 160–162, 170].

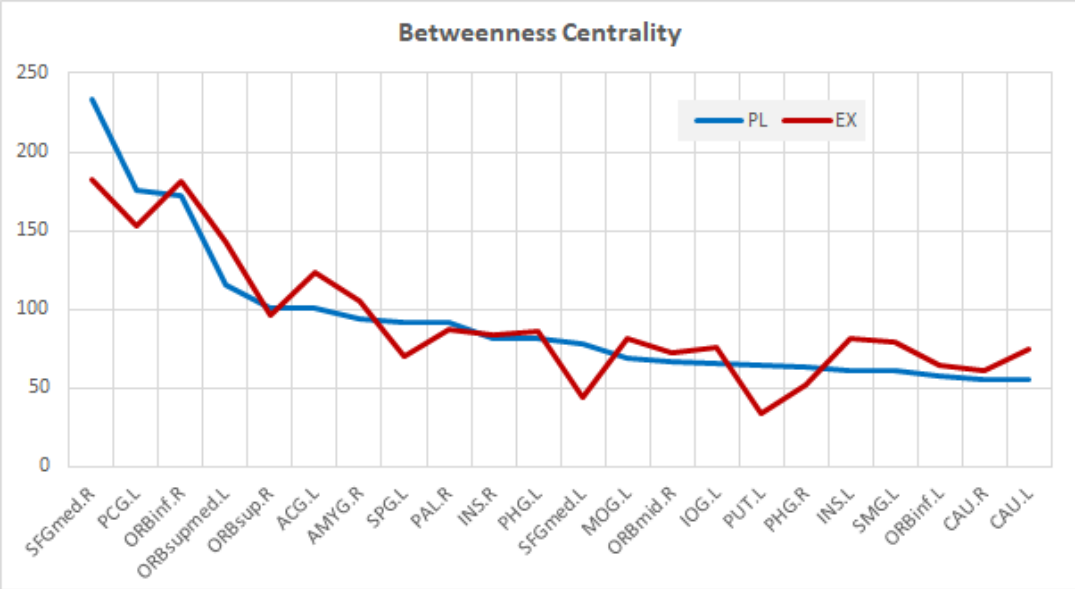


Figure 5.31: Nodes with high betweenness centrality for planning and execution tasks.

5.4 Classification of Planning and Execution tasks of Complex Problem Solving

Until now, we used an information-theoretic method for modeling the brain connectome. Measurements of information interchange across the anatomic regions are crucial since their purpose is to accurately draw conclusions that were not previously known about the data in question. In this section, we test the validity of the suggested dynamic brain networks by training a classifier with the arc-weights of the estimated dynamic brain networks and measuring the test performances of the planning and execution tasks.

We applied various classification models on TOL fMRI data, for example decision trees, discriminant analysis, Support Vector Machines, logistic regression, nearest neighbors, and ensemble classification. Support vector machines (SVM) models give very good results for the classification of TOL tasks.

5.4.1 Planning and Execution Tasks Classification Using Entropy and BOLD Values

In the first set of experiments, we provide the classification performances of two baseline methods: In the first method, we use the representative time series (BOLD values). In the second method, we estimate the first order entropy, obtained from each anatomical region, as explained in Chapter 4. Then fed the entropies of anatomical regions into a Support Vector Machine classifier. In both methods, we measure the

Table 5.2: Classification performances of SVM for planning and execution tasks

Performance (%)	**BOLD**	**ENTROPY**	**PEARSON**	**K-L**
All sessions	91.10	91.46	90.98	91.57
Successful sess.	91.62	91.54	90.52	92.31
Unsuccessful sess.	90.87	91.22	90.61	91.47

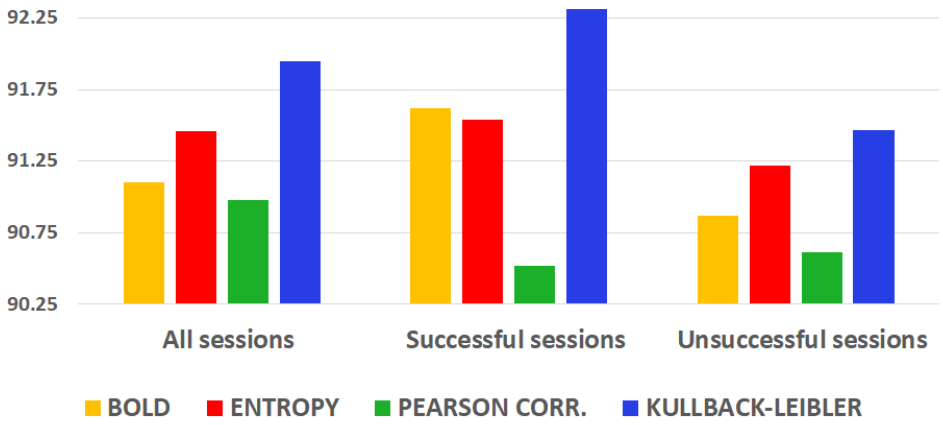


Figure 5.32: Classification performances (%) of SVM classifier for planning and execution tasks.

classification performance of planning and execution phases, using the labeled fMRI data.

In the baseline all of the experiments we split the data into successful and unsuccessful sessions. The successful 42 sessions consist of 8.192 representative time series of anatomic regions. We make another test for unsuccessful sessions (30 sessions), that have total 6.757 representative time series for each anatomic region. Finally, we use the data from all Sessions (72 sessions), total of 14.949 records. The classification performances of planning and execution tasks for BOLD fMRI time series, as representative time series of anatomical regions and region entropy are given in Table 5.2 (left and middle left columns), at the output of Support Vector Machine algorithm, with Gaussian kernel. Within-subject performances of SVM classifier for planning and execution tasks is given in Table 5.3 (left and middle columns) and Figure 5.33.

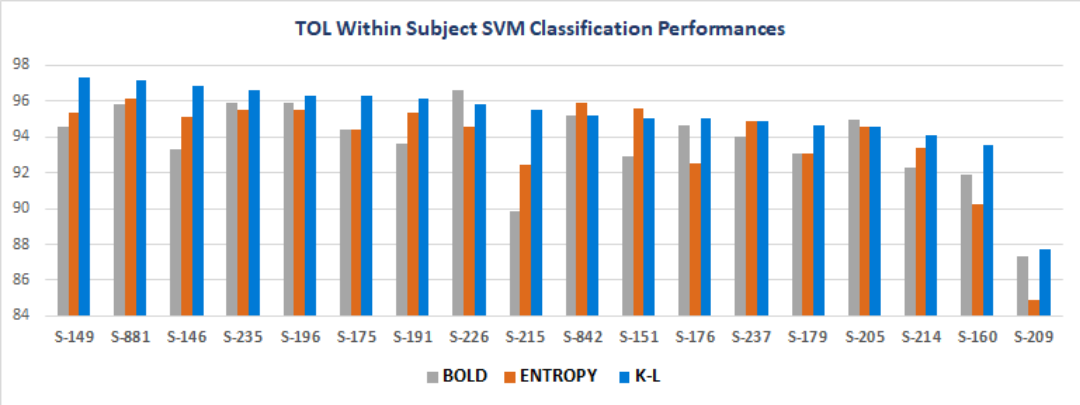


Figure 5.33: Within-subject classification performance (%) of SVM classifier for planning and execution tasks.

5.4.2 Planning and Execution Tasks Classification Using Kullback-Leibler and Pearson Correlation Matrices

In this set of experiments, we use Kullback-Leibler adjacency matrices in a vector form. We also used Pearson correlation matrices in vector form at the input of an SVM classifier to discriminate the planning and execution phases.

In the first test, data from all subjects are used for classification. Then, we test the classification performances for successful and unsuccessful sessions, consecutively. The planning and execution time instances are selected from voxel based TOL fMRI data. Then, dynamic K-L matrices and Pearson correlation matrices are generated for each time instance of the planning and execution phases.

In order to reduce the curse of dimensionality problem, we reduced the size of the input space by discarding the high-entropy anatomic regions. By using cross validation technique, we select the K-L divergence across the 60 low entropy anatomic regions to train an SVM classifier, with linear kernel. The classification results are given in the middle right and the right most columns of Table 5.2. Figure 5.32 shows the comparison of classification performances of BOLD, Entropy, Pearson Correlation, and Kullback-Leibler methods for planning and execution tasks.

The within-subject performances of SVM classifier for planning and execution tasks is given in Table 5.3 and Figure 5.33. In table 5.3, subjects that successfully solve the TOL problems are indicated in blue, and those that are unsuccessful are indicated in dark red.

Table 5.3: Within-Subject Performances for Planning and Execution Tasks

Subjects	BOLD	Entropy	Kullback-Leibler
S-146	93.33	95.11	**96.89**
S-149	94.57	95.35	**97.29**
S-151	92.95	**95.61**	95.02
S-160	91.90	90.28	**93.52**
S-175	94.44	94.44	**96.30**
S-176	94.61	92.53	**95.02**
S-179	93.10	93.10	**94.64**
S-191	93.62	95.32	**96.17**
S-196	95.92	95.51	**96.33**
S-205	**94.96**	94.54	94.54
S-209	87.35	84.90	**87.76**
S-214	92.25	93.36	**94.10**
S-215	89.85	92.48	**95.49**
S-226	**96.64**	94.54	95.80
S-235	95.91	95.54	**96.65**
S-237	93.99	**94.85**	**94.85**
S-842	95.19	**95.93**	95.19
S-881	95.79	96.14	**97.19**
Avg.	93.69	93.86	**95.15**

5.5 Summary and Interpretation of Results

In this chapter, we provided our complex problem solving experiments and data analysis results. First, we explain the Tower of London problem-solving test, our experimental setup, fMRI data collection environment, and the structure of the data.

For the localization of the activated anatomical regions, we use Shannon information entropy measurement. Our assumption is the low entropy regions are more intimately involved in complex problem-solving processes compared to the high entropy regions. Therefore, we investigate the relationship between the problem-solving task phases and the entropy measures. We estimate the information content of brain regions by Shannon entropy, at each anatomical region for two main phases of problem-solving and for expert and novice players.

We present two new definitions of entropy: Static entropy, and Dynamic entropy. Static entropy measures the entropy of each region, independent of time. It is defined over a period of time while the subject is in a resting state or performing a cognitive task. Dynamic entropy is used to analyze the change of the information content of brain regions at each time instant during the resting-state and the cognitive task. Both entropy measures enable us to study and compare the information content of anatomic regions for the resting states and the phases of complex problem solving.

We investigate the Dynamic entropy in two sets of experiments. In the first set, we investigate the dynamic entropy variations with respect to time for high and low entropy

regions during the resting-state and problem-solving task. In order to compare the behavior of the dynamic entropy in the anatomical regions, we estimate the dynamic entropy for each subject, session, time instance, and anatomical region. We average all the dynamic entropy measures over 42 successful sessions across anatomic region and time instances.

We observe that Entropy variations with respect to time are more structured for low-entropy anatomic regions, whereas randomly fluctuate for high-entropy regions. The dynamic entropy in the region is always low when the subject plays the TOL game, and it is always high when the subject is in a resting state. This indicates that playing the TOL game generates a more organized signal in a low entropy region. In other words, an anatomical region that produces a more orderly outcome has a lower entropy value. Remember that a low entropy value indicates greater predictability of the region behavior. On the other hand, when we analyze dynamic entropy variations in a high entropy region, we observe that dynamic entropy shows a more irregular course of action, appearing to be highly contaminated by noise.

In the second set of experiments, we study the behavior of dynamic entropy changes for the planning and execution phases of complex problem-solving, and resting states. We analyze the behavior of dynamic entropy in the low entropy anatomical regions of expert and novice players. The dynamic entropy measures show that for the low entropy regions there is a substantial difference between the expert and novice players. For the expert player, the dynamic entropy in the resting state and planning-execution phases are highly correlated with the experimental setup. The highest information content is observed during the resting state in a low entropy region, indicating that the region of interest can be in many states. A particular state is unpredictable. The information content is decreased in the region activated by the problem-solving tasks. This increases the probability that the relevant region is in a certain state. However, for the novice player, the dynamic entropy and task relations are more random compared to the expert players. The dynamic entropy fluctuations are rather arbitrary during the resting state and while playing the TOL game. This indicates that, the probability of problem-solving related regions being in a particular state is very low.

We analyze the low dynamic entropy regions during the resting states between two puzzles in the TOL game. We compare low entropy regions and several resting-state networks and pointed out similarities.

We estimate the static entropy for planning and execution tasks in two sets of experiments. First, we use all the voxels to estimate the average BOLD values for each anatomical region separately, and, then, estimate the static entropy. We observe that static entropy for the planning phase is relatively small compared to the execution. Second, we estimate static entropy values based on the most informative voxels selected by the ANOVA technique. We find the representative voxel time series of each anatomical region by using just the most informative voxels, then estimate static entropy. We see that the difference between static entropy measures of planning and execution phases is more accentuated when we eliminate the irrelevant voxels.

We observe that the difference between planning and execution entropy is higher for successful runs than the unsuccessful runs. This can be interpreted as the underlying anatomical regions of an expert player is more organized and predictable compared to that of a novice player.

Researches on problem-solving have shown that the expert and novice players think and solve problems in different ways [177, 178, 183]. When we compare the low entropy regions, we observe that experts have lower static entropy values compared to novices. Also, the durations for planning and execution phases across the successful and unsuccessful sessions show that the experts spent more time on planning than execution, whereas novices spent less time for planning, more time on the plan execution. It is suggested that, if the planning is successful, the execution task will be completed in a relatively short time, otherwise re-planning is required during execution to make corrections on the preliminary plan.

Another important goal of this study is to represent brain activities by network structure which can model the cognitive states of complex problem solving for successful and unsuccessful problem-solving sessions. We propose a novel method to generate static and dynamic brain networks using Kulback-Leibler divergence, which is also called relative entropy.

We assume that the degree of co-activation between two anatomic regions can be measured by Kullback-Leibler divergence. Therefore, the measure of K-L divergence between the anatomic regions is used as the arc weights of the brain network formed among the anatomical regions.

We estimate two types of brain networks, namely, static and dynamic brain networks. First, dynamic brain networks are estimated from the distribution of voxel intensity values for each anatomic region, at each time instant. The Kullback-Leibler divergence between the anatomic regions is then estimated for the distribution functions of the region pairs and a connectivity graph at each time instant is produced. The connection pairs are processed according to the lowest K-L distances and highest occurrences to find out the significant connections. Since we obtain a probability distribution function for each time instant, over all the voxels residing in a region, it is assumed that this approach gives a dynamic nature of brain connectivity between anatomical regions.

We estimated the Pearson correlation coefficients as a baseline method for the same distributions. Although the K-L distance and Pearson R values are different metrics, the two matrices display similar patterns.

Then, we estimate the static brain networks to investigate the behavior of the planning and execution tasks for successful and unsuccessful problem-solving sessions. Static brain networks are estimated by taking the expected value over all time instances to represent the K-L divergences for all the planning and execution phases. We represent the planning and execution phases of each subject by the estimated node adjacency matrix of the brain network.

We observe that, for most cases, the number of minimum K-L distances between anatomic regions for the planning phase, is higher than the number of minimum K-L distances for the execution phase. This result is consistent with experimental neuroscience, which observes that planning is a more complicated cognitive process than the execution of the plan.

We test the validity of the dynamic K-L brain networks by training a classifier with the arc-weights of the estimated brain networks. We measure the classification performances for the planning and execution tasks using BOLD values, Entropy values, Pearson correlation coefficients, and Kullback-Leibler distances in a range of setups. By using the cross-validation technique, we select the Kullback-Leibler divergence across the 60 low entropy anatomic regions to train an SVM classifier, with the Linear kernel. The Kullback-Leibler gives the best performance, which has more than 90% classification accuracy. This result shows the representation power of our brain network model.

CHAPTER 6

CONCLUSION AND FUTURE DIRECTIONS

> "I thought becoming myself was improving each part piece by piece. But it was finding a hidden wholeness seeing the fractures as the design."
>
> Brianna Wiest

Our thoughts, feelings, and behaviors are rooted in our brains. How the physical processes in our brains transform into our thoughts, behaviors and perceptions is still not fully explained. This first requires explaining how to map the physical activity of the brain with its mental activity. Identifying the parts of the brain that perform a particular mental task and the functional brain networks activated by this process is crucial to understanding this issue.

In this thesis, we develop a computational model based on the Shannon information theory, for localizing the brain regions that contribute to the execution of complex problem solving and constructing underlying dynamic and static brain network representations. Using this model, we are able to represent the phases of complex problem solving, which can be vaguely observed by simple statistical techniques. This also allows to understand whether the phases of complex problem-solving as hypothesized by theory [10] are present and detectable in fMRI data.

In this context, we first detect anatomical regions that become active during solving complex problems. Second, we build static and dynamic brain networks that identify and model the main phases of complex problem solving. Our method provides a basis for comparison of the common characteristics and activation differences between individuals during problem solving by detecting activated brain regions and brain networks for strong and weak problem solvers. The proposed model can successfully distinguish the differences in the brain activity of expert and novice problem solvers for the planning and execution phases of complex problem solving using fMRI data.

We showed that estimating the entropy of brain regions and the relative entropy between those regions provide an informative tool for understanding brain states in the process of solving complex problems. The proposed entropy measures effectively identify the active brain regions which participate in the complex problem solving task, where the entropy values are relatively low.

We observe that dynamic entropy variations are more structured for low-entropy regions, whereas they fluctuate randomly for high-entropy regions. In a low entropy anatomic region, the resting state entropies are very high compared to the problem-solving states. We also show that active regions have relatively lower static entropy values for the planning task, compared to the execution task. Finally, while the dynamic entropy measures are very structured for expert players, they are more noisy for novice players. Expert players have lower static entropy values compared to novice players.

Comparison of our results with the experimental studies in the literature indicates a high overlap between the low entropy regions and active anatomic regions observed by experimental neuroscience, for the TOL game. Therefore, we can deduce that the regions with low entropy correspond to the active anatomic regions of TOL game playing.

We observe that relative entropy (Kullback-Leibler divergence) provides an information theoretic tool to estimate static and dynamic brain networks. The validity of the proposed network model is shown by the classification performance of the phases of complex problem solving. Our results show strong promise in using the Kullback-Leibler networks as a measure for characterizing the brain states for a cognitive task.

Therefore, we conclude that the proposed computational models which is based on information entropy and relative entropy are capable of identifying active regions, detecting brain states and predicting brain networks associated with a specific cognitive process.

Entropy can be helpful as a bridging tool that would enable a better understanding of the mind-brain relationship and the mechanics of cognition [184]. Understanding the neural basis of a particular cognitive task is expected to provide new insights into the mechanisms of internal information representation and processing behind mental processes. We suggest that our study can lead to a better characterization of processes that support complex problem solving.

6.1 Critiques of the Proposed Method

The major strengths of the proposed model might be its reliability, flexibility, and generality. It can provide a simple and useful heuristic, enabling the translations to be made between a quantitative measure of the brain activity and the cognitive processes.

The reliability depends on the fact that the model is based directly on a widely accepted and validated theory that provides a framework and principled approach to implementation. A robust theory brings its explanatory power for the correct interpretation of the results. Besides, the data obtained as a result of the implementation of our model can feed machine learning algorithms to test the validity of the model.

The proposed model can be applied to most of the neuroscientific data types. Also, any directed or undirected brain connectivity graph can be created using the proposed

model. It facilitates simple comparisons between voxels, anatomical regions, cognitive tasks, or subjects. Furthermore, it can quantify representational interactions between different neuroimaging responses, such as different regions, frequency bands, time periods, or recording modalities. Information about different brain signals, such as, spikes and local field potentials, can be analyzed by Shannon entropy, which allows a comparison between them as they map the two signals on a common scale. [139].

However, there are some conceptual and methodological issues regarding our method that we would like to discuss here.

- The issue of whether cognitive functions are localized in certain brain regions and whether there is a one-to-one match between brain areas and cognitive processes has been debated for many years. The widely accepted view is that cognitive tasks are performed not by isolated brain regions, but they are functionally connected with each other. Several distributed brain areas are orchestrated in the performance of even simple cognitive tasks.

- It is known that fMRI data analysis method has some limitations. The basic shortcomings of the fMRI analysis were explained in the "Conceptual and Methodological Problems" section. It is accepted that fMRI research on localization provides valuable information for understanding how the brain and mind works. However, as the other neuroimaging modalities (i.e., EEG, MEG), it will not provide us with complete information about how the brain represents, encodes, and executes cognitive processes.

- Information theory does not provide information about directionality or causality. Although we did not implement the causality in our method, we achieved directionality by estimating Kullback-Leibler networks.

  Various techniques from information theory were used for the detection of causal relationships in multivariate systems [185]. In particular, it is shown that Conditional Mutual Information, and Transfer Entropy [186] can detect and measure causal relation and information flow between observed variables. These techniques are the nonlinear extensions of the Granger causality. They are used in numerous applications in various fields of sciences, such as Palus et al. [187], Hinrichs et al. [188], Pflieger and Greenblatt [189], Vicente et al. [190]. The partial transfer entropy (PTE) [191] is recently introduced as the extension of transfer entropy to account for the presence of confounding variables. These two techniques are actually the expressions of the conditional mutual information.

  In several applications, directed information has been shown to characterize statistically causal effects between random processes in a more robust and meaningful way than Granger causality. Quinn et al [192] connect the information-theoretic concept of "directed information" to Granger's philosophical relationship between causality and prediction. They show that an important difference between directed information and Granger causality is that, directed information is a sum of K-L divergences. Thus, is well-defined for arbitrary

joint probability distributions. They also present that directed information is more flexible as a metric that can be directly applicable to many modalities, including neural spike trains. Waddell et al. [193] proposed causal entropy, a time-varying measure of the regularity of the firing relationship of one neuron with respect to another neuron over a period of time. They used an entropy measure to monitor changes in functional, temporal interdependence between two neurons. Amblard and Michel's [194] study focused on adapted tools for investigating Granger causal relationships, which are the conditional transfer entropy rate and the conditional instantaneous information change rate.

- The available literature on problem solving task suggest that there are two basic phases, called planning and execution, as a sequential processes. However, recent studies suggest that the interaction between these processes can be quite dynamic depending on the characteristics of the problem [147].

6.2 Future Work

So far, we presented how the information-theoretic framework can be effectively used in fMRI data analysis. We now suggest some future work to improve and extend this method for the successful use of information theory in the study of cognition with its quantitative tools.

- In this study, we focused on understanding complex problem solving by examining its neural underpinnings. Information theory allows us to quantify information processing in various domains of cognitive processing. Testing the proposed model on other cognitive tasks is crucial to demonstrate the robustness of the model and to better understand these processes.

- As mentioned in the previous section, there are several approaches to quantify causality between multivariate time series under the framework of information theory. Extending our proposed model to include causal inference will be of great benefit in explaining cognitive processes more effectively. In this respect, Conditional Mutual Information and Transfer Entropy would be the most prominent ones. However, transfer entropy has some lacks which lead to high computational complexity and redundancy [195, 196]. An improved version of Transfer Entropy making use of mutual information-driven state space reconstruction technique for time series analysis was developed, namely mutual information on mixed embedding (MIME) [195]. Further, it improved to partial MIME (PMIME) [196]. It is suggested that both methods are computationally efficient and have wide applications to various data analyses [197].

- The proposed model can be applied to simultaneously multimodal recorded data of the neural activity, such as EEG. For example, simultaneous EEG-fMRI offers a neuroimaging method that combines the advantages of both modalities. In this context, the information-theoretic approach provides a clear framework that allows questions of functional specialization and functional integration to

be addressed in a unified framework. It may also let us to investigate the key questions regarding possible synergistic effects on a model-free basis [98]. The synergistic effects between the two modalities are important for linking the different selectivity of EEG and fMRI to different stimulus features. Thus, for the representation of the different neurobiological substrates underlying each imaging modality. As a result of such an approach, analysis of combined EEG-fMRI recording data would allow for a more comprehensive characterization of neuronal processes under investigation.

- Many studies have tested and evaluated various types of entropy measures on fMRI data. Different variations of Shannon entropy is adapted by the studies, including differential entropy, permutation entropy, multiscale entropy, multiscale permutation entropy, and sample entropy [85–89]. Although there are a variety of entropy measurements, most of them are computationally similar. However, it will be useful for future research to further examine the efficiency of various entropy measures.

- Estimating entropy from fMRI data is a challenging task. It can be said that the most critical issue in entropy estimation is estimating the probability distribution accurately. We applied a non-parametric representation of the probability density function, namely Kernel Density Estimation. Accordingly, we select proper kernel function and bandwidth depending on the properties of the data. Recently, theoretical and empirical evidence suggests that lognormal (skewed) distributions are suitable to biological signals. It may be useful to examine the lognormal distribution for entropy estimation.

  The rationale behind this suggestion can be explained as follows. Neural systems involve many interactive processes and the multiplication of many variables. The outcome of this multiplication process is mostly a skewed distributions that can be described by logarithmic rules. It is shown that distance perception and time perception vary logarithmically with the distance length and the time interval; word usage and sentence lengths in most languages follow a lognormal form. The mental organization of numbers is also described by a logarithmic scale [120]. It is reported that numerous properties of neuronal activity, such as the average firing rates of neurons, the rate and magnitude of spikes, the magnitude of population synchrony, and the correlations between presynaptic and postsynaptic spikes show approximate lognormal distributions [198]. It is observed that the logarithmic scale distribution of weights and excitability appears to be a general, functional property of the neural system [199].

  Since a logarithmic transformation brings log-normal data back to normal and converts multiplication to addition, it is suggested that the log-normal distribution can be selected as a worthy choice when “amounts” are modeled – these being quantities that can only take positive values [200]].

- Another area of future research is the relationship between entropy and cognitive disorders. Cognition enables us to perceive and connect to the world. Any disruption of this ability can have serious consequences. Some studies have already been done using information theoretic methods to understand the origin

of these deficiencies [201,202]. The dynamic and static entropy measurements, and Kulback-Leibler brain networks, proposed in this study, can be used as a potential marker of brain diseases. The changes in the information content of anatomical regions and Kullback-Leibler functional brain networks can be investigated in neurological diseases and disorders. These findings can also be used for diagnosis of neurological diseases.

- One of the hitherto unsolved questions is how our mind uses information provided by external or internal phenomena to reduce uncertainty about a particular situation. Then, it decide the appropriate behavior given the most probable state of the world. It constantly shapes its view by making predictions about the current and future state of the world. Information theory provides tools for measuring probability relationships and uncertainty. Although utility maximization or reward motivation has traditionally been seen as the driving force behind the choice behavior, it has recently been suggested that the minimization of uncertainty may exceed the reward motivation. Therefore, it might be an essential drive for the brain [203].

  Behavior analysis has begun to adopt the information theoretic tools as a new way of measuring the interrelationships between behavior, stimuli, and possible outcomes [15]. Studies investigating this topic propose that informational measures appear to be a reasonable basis for predicting behavior.

  Behavioral studies investigating the probabilistic mechanism of the brain have typically focused on Bayesian models. It is argued that integration of Bayesian and information theoretic frameworks is inherently compatible. It is also claimed that the time average of Bayesian model evidence is the same as the entropy of sensory data sampled by a brain [204]. This means that a Bayesian brain trying to maximize its evidence is implicitly trying to minimize its entropy. Thus, it provides a principled approach for self-organization against the natural tendency for the disorder.

  It is known that Shannon's information definition is associated with the actual symbols (letters, numbers, etc.) which are used to encode a message. It is not intended to relate to the meaning of the messages that they transmit. Thus, it is independent of semantic meaning; it depends on the presentation probabilities [205]. However, considering Bayesian approaches to explain brain functioning, it is suggested that information theory has the basic infrastructure for making semantic inferences.

  Interpreting information theory in the light of the Bayesian brain hypothesis may allow us to use formal structures of information theory to describe behavior in terms of beliefs and inferences. If events in a given domain differ in their probability measure, information uncertainty may be a valid alternative explanation that does not rely on domain-specific semantic explanations [205]. Behavioral theories using this hybrid approach would have the potential to provide a solid conceptual and mathematical framework for understanding the changing representations of the environment that stimulate behavior.

  Information theory can provide a common currency for relating external and internal phenomena to behavior. It can also serve as a common language between

disciplines that would otherwise remain divided by independent theoretical and methodological trajectories [15]. The integrative potential of information theory would enable isolated behavior analysis approaches to establish a common ground in this framework and build bridges to other areas of brain-mind studies.

REFERENCES

[1] C. Eliasmith, "How to build a brain - a neural architecture for biological cognition," *Oxford University Press, ISBN 978–0–19–979454–6*, 2013.

[2] F. M. Al-Shargie, "Early detection of mental stress using advanced neuroimaging and artificial intelligence," *arXiv:1903.08511 [q-bio.NC]*, 2019.

[3] M. Soltanlou, M. A. Sitnikova, H. C. Nuerk, and T. Dresler, "Applications of functional near-infrared spectroscopy (fnirs) in studying cognitive development: The case of mathematics and language," *Front Psychol. 2018;9:277. Published 2018 Apr 3. doi:10.3389/fpsyg.2018.00277*, 2018.

[4] J. Richiardi, S. Achard, H. Bunke, and D. V. D. Ville, "Machine learning with brain graphs: predictive modeling approaches for functional imaging in systems neuroscience," *IEEE Signal Processing Magazine: 58–70*, 2013.

[5] C. E. Shannon, "A mathematical theory of communication," *The Bell System Technical Journal, Vol. 27, pp. 379–423, 623–656*, 1948.

[6] I. G. et al., "Blunted prefrontal perfusion in depressed patients performing the tower of london task," *Psychiatry research. 139. 31-40.10.1016/j.pscychresns.2004.09.007*, 2005.

[7] W. contributors, "Kernel density estimation," *Wikipedia, the free encyclopedia. URL https://en.wikipedia.org/wiki/Kernel_density_estimation*, 2019.

[8] R. G. Osuna, "Pattern analysis," *CSCE-666 Pattern Analysis. Dept of Computer Science and Engineering, Texas A&M University*, 2020.

[9] B. K. et al., "Methodological problems on the way to integrative human neuroscience," *Front Integr Neurosci. 10:41. doi: 10.3389/fnint.2016.00041. PMID: 27965548; PMCID: PMC5126073*, 2016.

[10] H. A. Simon and A. Newell, "Human problem solving: The state of the theory in 1970," *American Psychologist, 26(2), 145–159*, 1971.

[11] Y. Wang and V. Chiew, "On the cognitive process of human problem solving," *Cognitive Systems Research: An International Journal, Elsevier, 11(1), 81-92*, 2010.

[12] S. Keshmiri, "Entropy and the brain: An overview," *Entropy 2020, 22(9), 917*, 2020.

[13] W. M. Haschek, C. G. Rousseaux, and M. A. Wallig, "Nervous system," *Fundamentals of Toxicologic Pathology (Second Edition), Academic Press,2010, Pages 377-410, ISBN 9780123704696*, 2010.

[14] M. Lindquist, "The statistical analysis of fmri data," *Statistical Science, 23, 439-464*, 2008.

[15] G. Jensen, R. D. Ward, and P. D. Balsam, "Information: theory, brain, and behavior," *J Exp Anal Behav. 2013;100(3):408-431. doi:10.1002/jeab.49*, 2013.

[16] L. Zyga, "Phase transition may explain how brain neurons encode information," *Phys.org report*, 2013.

[17] C. Koch and G. Marcus, "How does the brain speak to itself?. cracking the brain's codes," *MIT Technology Review*, 2014.

[18] B. J. Richmond, "Information coding," *Encyclopedia of Neuroscience, Academic Press, Pages 137-144, ISBN 9780080450469*, 2009.

[19] B. He, "Neural engineering," *Springer, ISBN: 978-3-030-43395-6 https://doi.org/10.1007/978-3-030-43395-6*, 2020.

[20] T. Gollisch and M. Meister, "Rapid neural coding in the retina with relative spike latencies," *Science. 319(5866): 1108–1111. doi:10.1126/science.1149639. ISSN 0036-8075*, 2008.

[21] D. A. Butts, C. Weng, and J. J. et al., "Temporal precision in the neural code and the timescales of natural vision," *Nature. 449 (7158): 92–5. doi:10.1038/nature06105. PMID 17805296*, 2007.

[22] J. M. Pearce, "Marie-jean-pierre flourens (1794-1867) and cortical localization," *Eur Neurol. 2009;61(5):311-4. doi: 10.1159/000206858. PMID: 19295220*, 2009.

[23] S. Clare, "Functional magnetic resonance imaging: Methods and applications," *Ph.D. Thesis, The University of Nottingham*, 1997.

[24] S. Clare, "Functional mri : Methods and applications," *University of Nottingham, PhD thesis*, 1997.

[25] B. Vallabhajosula, "The basics of neuroimaging," *Print ISBN-13: 9780199995721. Published to Oxford Scholarship Online. DOI:10.1093/acprof:oso/9780199995721.003.0003*, 2015.

[26] K. Rajamanickam, "A mini review on different methods of functional-mri data analysis," *Arch Intern Med Res 2020; 3(1): 044-060*, 2020.

[27] G. Kalantar, "Advanced signal processing solutions for brain-computer interfaces: From theory to practice," *MS Thesis, The Department of Information Systems Engineering Concordia University Montreal, Canada*, 2018.

[28] S. Williams and R. Karim, "Physiological psychology," *Notion Press, eISBN 978-1-64324-266-8*, 2018.

[29] M. E. Ladd, P. Bachert, M. Meyerspeer, E. Moser, A. M. Nagel, and D. G. N. et al., "Pros and cons of ultra-high-field mri/mrs for human application," *Progress in nuclear magnetic resonance spectroscopy. 109: 1-50. doi:10.1016/j.pnmrs.2018.06.001 - Pubmed*, 2018.

[30] J. R. P. et al., “Fast fmri can detect oscillatory neural activity in humans,” *PNAS October 25, 2016 113 (43) E6679-E6685*, 2016.

[31] K. Cultice, “Handbook of eeg interpretation,” *The Neurodiagnostic Journal 47 (4), 344*, 2007.

[32] D. Bente, P. Chenchanna, W. Scheuler, and P. Sponagel, “Drug-induced changes of eeg vigilance and optimizing control behavior during car driving,” *EEG-EMG-Zeitschrift Fur Elektroenzephalographie Elektromyographie Und Verwandte Gebiete 9 (2), 61-73*, 1978.

[33] S. Bunge and I. Kahn, “Cognition: An overview of neuroimaging techniques,” *Encyclopedia of Neuroscience, vol. 2, pp. 1063-1067*, 2009.

[34] F. Cazals and P. Kornprobst, “Modeling in computational biology and biomedicine - a multidisciplinary endeavor,” *Springer ISBN 978-3-642-31207-6; ISBN 978-3-642-31208-3 (eBook); DOI 10.1007/978-3-642-31208-3*, 2013.

[35] A. Villringer and B. Chance, “Non-invasive optical spectroscopy and imaging of human brain function,” *Trends in Neurosciences 20 (10), 435-442*, 1997.

[36] M. Ferrari and V. Quaresima, “A brief review on the history of human functional near-infrared spectroscopy (fnirs) development and fields of application,” *NeuroImage 63 (2), 921-935*, 2012.

[37] A. Wright, “Brain scanning techniques (ct, mri, fmri, pet, spect, dti, dot),” *Cerebra*, 2010.

[38] E. W. Lang, A. M. Tome, I. R. Keck, J. M. Gorriz-Saez, and C. G. Puntonet, “Brain connectivity analysis: a short survey,” *Comput Intell Neurosci. 2012:412512. doi:10.1155/2012/412512*, 2012.

[39] J. A. Etzel, J. M. Zacks, and T. S. Braver, “Searchlight analysis: Promise, pitfalls, and potential,” *NeuroImage, Volume 78, 2013, Pages 261-269, ISSN 1053-8119*, 2013.

[40] S. Sarraf and S. Jian, “Functional brain imaging: A comprehensive survey,” *arXiv:1602.02225v4 Medical Physics*, 2016.

[41] A. Mahmoudi, S. Takerkart, F. Regragui, D. Boussaoud, and A. Brovelli, “Multivoxel pattern analysis for fmri data: a review,” *Comput Math Methods Med. 2012;2012:961257. doi:10.1155/2012/961257*, 2012.

[42] J. V. Haxby, M. I. Gobbini, M. L. Furey, A. Ishai, J. L. Schouten, and P. Pietrini, “Distributed and overlapping representations of faces and objects in ventral temporal cortex,” *Science. 2001 Sep 28; 293(5539):2425-30*, 2001.

[43] D. D. Cox and R. L. Savoy, “Functional magnetic resonance imaging (fmri) "brain reading": detecting and classifying distributed patterns of fmri activity in human visual cortex,” *Neuroimage. 2003 Jun; 19(2 Pt 1):261-70.*, 2003.

[44] A. J. O'Toole, F. Jiang, H. Abdi, and J. V. Haxby, "Partially distributed representations of objects and faces in ventral temporal cortex," *J Cogn Neurosci. 2005 Apr; 17(4):580-90*, 2005.

[45] M. Spiridon and N. Kanwisher, "How distributed is visual category information in human occipito-temporal cortex? an fmri study," *Neuron. 2002 Sep 12; 35(6):1157-65*, 2002.

[46] K. A. Norman, S. M. Polyn, G. J. Detre, and J. V. Haxby, "Beyond mind-reading: multi-voxel pattern analysis of fmri data," *Trends in Cognitive Sciences, Volume 10, Issue 9, 2006, Pages 424-430, ISSN 1364-6613, https://doi.org/10.1016/j.tics.2006.07.005*, 2006.

[47] E. Formisano, F. D. Martino, and G. Valente, "Multivariate analysis of fmri time series: classification and regression of brain responses using machine learning," *Magn Reson Imaging. 2008 Sep;26(7):921-34*, 2008.

[48] B. B. Biswal and F. Z. Y. et al., "Functional connectivity in the motor cortex of resting human brain using echo-planar mri," *Magn. Reson. Med. 34, 537–541*, 1995.

[49] A. M. Bastos and J. M. Schoffelen, "A tutorial review of functional connectivity analysis methods and their interpretational pitfalls," *Front Syst Neurosci. 2016 Jan 8;9:175*, 2015.

[50] Y. W. et al., "An efficient and reliable statistical method for estimating functional connectivity in large scale brain networks using partial correlation," *Frontiers in Neuroscience, Volume:10*, 2016.

[51] F. T. Sun, L. M. Miller, and M. D'Esposito, "Measuring interregional functional connectivity using coherence and partial coherence analyses of fmri data," *Neuroimage 21, 647–658*, 2004.

[52] F. G. Ashby and J. G. Waldschmidt, "Fitting computational models to fmri data," *Behavior Research Methods 2008, 40 (3), 713-721*, 2008.

[53] K. J. Friston, "Statistical parametric mapping and other analyses of functional imaging data," *Brain Mapping: the Methods. Academic Press, Toga, A.W., Mazziotta, J.C. (Eds.), San Diego, CA, pp. 363– 396*, 1996.

[54] Y. Zang, T. Jiang, Y. Lu, Y. He, and L. Tian, "Regional homogeneity approach to fmri data analysis," *Neuroimage. 2004 May;22(1):394-400*, 2004.

[55] L. Kaiming, G. Lei, N. Jingxin, L. Gang, and L. Tianming, "Review of methods for functional brain connectivity detection using fmri," *Computerized Medical Imaging and Graphics, Volume 33, Issue 2, 2009, Pages 131-139*, 2009.

[56] J. E. Jackson, "A user's guide to principal components," *New York, NY: Wiley*, 1991.

[57] A. Hyvarinen, "Independent component analysis: recent advances," *Philosophical Transactions: Mathematical, Physical and Engineering Sciences. 371 (1984): 20110534*, 1984.

[58] F. V. Farahani and W. Karwowski, "Computational methods for analyzing functional and effective bran network connectivity using fmri," *Advances in Neuroergonomics and Cognitive Engineering: Proceedings of the AHFE 2018 Int. Conf. on Neuroergonomics and Cognitive Engineering, Florida USA*, 2018.

[59] F. G. Ashby, "Statistical analysis of fmri data," *Cambridge, Mass. : MIT Press, ISBN: 9780262295697, 352 pp., March 2011*, 2011.

[60] A. Hyvarinen, "Independent component analysis: recent advances," *Philos Trans A Math Phys Eng Sci. 2012;371(1984):20110534. Published 2012 Dec 31. doi:10.1098/rsta.2011.0534*, 2012.

[61] C. Windischberger, M. Barth, C. Lamm, L. Schroeder, H. Bauer, R. C. Gur, and E. Moser, "Fuzzy cluster analysis of high-field functional mri data," *Artif Intel Med. 2003;29(3):203–223*, 2003.

[62] A. Ben-Hur, D. Horn, H. Siegelmann, and V. Vapnik, "Support vector clustering," *The Journal of Machine Learning Research, vol. 2, pp. 125–137*, 2001.

[63] C. Cortes and V. Vapnik, "Support-vector networks," *Machine Learning, vol. 20, no. 3, pp. 273–297*, 1995.

[64] V. N. Vapnik, "The nature of statistical learning theory," *eBook, Springer, ISBN 978-1-4757-3264-1*, 1995.

[65] M. Timothy, D. Alok, and V. S. et al., "Support vector machine classification and characterization of age-related reorganization of functional brain networks," *NeuroImage, vol. 60, no. 1, pp. 601–613*, 2012.

[66] S. J. Hanson and Y. O. Halchenko, "Brain reading using full brain support vector machines for object recognition: there is no "face" identification area," *Neural Computation, vol. 20, no. 2, pp. 486–503*, 2008.

[67] X. Zhang, F. Tokoglu, M. Negishi, J. Arora, S. Winstanley, D. D. Spencer, and R. T. Constable, "Social network theory applied to resting-state fmri connectivity data in the identification of epilepsy networks with iterative feature selection," *J Neurosci Methods. 2011 Jul 15;199(1):129-39*, 2011.

[68] A. Khazaee, A. Ebrahimzadeh, and A. Babajani-Feremi, "Identifying patients with alzheimer's disease using resting-state fmri and graph theory," *Clin Neurophysiol. 2015 Nov;126(11):2132-41*, 2015.

[69] M. Ozay, I. Oztekin, U. Oztekin, and F. T. Y. Vural, "Mesh learning for classifying cognitive processes," *arXiv:1205.2382v3*, 2012.

[70] O. Firat, M. Ozay, I. Onal, I. Oztekin, and F. T. Y. Vural, "Functional mesh learning for pattern analysis of cognitive processes," *2013 IEEE 12th International Conference on Cognitive Informatics and Cognitive Computing, 2013, pp. 161-167*, 2013.

[71] I. O. Ertugrul, M. Ozay, and F. T. Y. Vural, "Hierarchical multi-resolution mesh networks for brain decoding," *Brain Imaging Behav. 2018 Aug;12(4):1067-1083*, 2018.

[72] M. Khosla, K. Jamison, G. H. Ngo, A. Kuceyeski, and M. R. Sabuncu, "Machine learning in resting-state fmri analysis," *Magnetic Resonance Imaging, Volume 64, 2019, Pages 101-121*, 2019.

[73] S. Vigneshwaran, B. S. Mahanand, and S. Suresh, "Using regional homogeneity from functional mri for diagnosis of asd among males," *International Joint Conference on Neural Networks (IJCNN) (Killarney)*, 2015.

[74] P. Patel, P. Aggarwal, and A. Gupta, "Classification of schizophrenia versus normal subjects using deep learning," *10th Indian Conference on Computer Vision, Graphics and Image Processing (ICVGIP) (Guwahati)*, 2016.

[75] R. J. Meszlenyi, K. Z. Buza, and Z. Vidnyanszky, "Resting state fmri functional connectivity-based classification using a convolutional neural network architecture," *Front. Neuroinformatics 11:61*, 2017.

[76] H. I. Suk, C. Y. Wee, S. W. Lee, and D. Shen, "State-space model with deep learning for functional dynamics estimation in resting-state fmri," *Neuroimage 129 (2016) 292-307*, 2016.

[77] X. Guo, K. C. Dominick, A. A. Minai, H. Li, C. A. Erickson, and L. J. Lu, "Diagnosing autism spectrum disorder from brain resting-state functional connectivity patterns using a deep neural network with a novel feature selection method," *Front Neurosci 11 (2017) 460*, 2017.

[78] A. S. Heinsfeld, A. R. Franco, R. C. Craddock, A. Buchweitz, and F. Meneguzzi, "Identification of autism spectrum disorder using deep learning and the abide dataset," *NeuroImage. Clinical, 17, 16–23*, 2017.

[79] J. Kim, V. D. Calhoun, E. Shim, and J. H. Lee, "Deep neural network with weight sparsity control and pre-training extracts hierarchical features and enhances classification performance: Evidence from whole-brain resting-state functional connectivity patterns of schizophrenia," *Neuroimage 124 (Pt A) (2016) 127-146*, 2016.

[80] D. Nie, H. Zhang, and E. Adeli, "3d deep learning for multi-modal imaging-guided survival time prediction of brain tumor patients," *International Conference on Medical Image Computing and Computer-Assisted Intervention (Istanbul) 9901, 212–220*, 2016.

[81] O. Firat, E. Aksan, and I. Oztekin, "Learning deep temporal representations for fmri brain decoding," *1st International Workshop on Medical Learning Meets Medical Imaging (MLMMI) (Lille)*, 2015.

[82] B. B. Kivilcim, I. O. Ertugrul, and F. T. Y. Vural, "Modeling brain networks with artificial neural networks," *Graphs in Biomedical Image Analysis and Integrating Medical Imaging and Non-Imaging Modalities, pp. 43–53. Springer*, 2018.

[83] C. Salvatore, P. Battista, and I. Castiglioni, "Frontiers for the early diagnosis of ad by means of mri brain imaging and support vector machines," *Curr. Alzheimer Res. 13, 509–533*, 2016.

[84] W. Yin, L. Li, and F. X. Wu, "Deep learning for brain disorder diagnosis based on fmri images," *Neurocomputing, ISSN 0925-2312*, 2020.

[85] W. Xiong, L. Faes, and P. C. Ivanov, "Entropy measures, entropy estimators, and their performance in quantifying complex dynamics: Effects of artifacts, nonstationarity, and long-range correlations," *Phys. Rev. E 2017, 95, 062114-1–062114-37*, 2017.

[86] B. Wang, Y. Niu, L. Miao, R. Cao, and P. Y. et al., "Decreased complexity in alzheimer's disease: Resting-state fmri evidence of brain entropy mapping," *Front Aging Neurosci. 9 378*, 2017.

[87] A. Davalos, M. Jabloun, P. Ravier, and O. Buttelli, "On the statistical properties of multiscale permutation entropy: Characterization of the estimator's variance," *Entropy 2019, 21, 450*, 2019.

[88] Z. Wang, Y. Li, A. R. Childress, and J. A. Detre, "Brain entropy mapping using fmri," *Plos One. 2014;9(3). pmid:24657999*, 2014.

[89] G. N. Saxe, D. Calderone, and L. J. Morales, "Brain entropy and human intelligence: A resting-state fmri study," *PLoS ONE 13(2): e0191582*, 2018.

[90] J. S. Richman and J. R. Moorman, "Physiological time-series analysis using approximate entropy and sample entropy," *American journal of physiology Heart and circulatory physiology 278: H2039–2049*, 2000.

[91] J. O. Maximo, C. M. Nelson, and R. K. Kana, "Unrest while resting? brain entropy in autism spectrum disorder," *Brain Res. 2021 Jul 1;1762:147435. PMID: 33753068*, 2021.

[92] Y. Jia, H. Gu, and Q. Luo, "Sample entropy reveals an age-related reduction in the complexity of dynamic brain," *Sci Rep 7, 7990*, 2017.

[93] N. LUdtke, N. K. Logothetis, and S. Panzeri, "Testing methodologies for the nonlinear analysis of causal relationships in neurovascular coupling," *Magn Reson Imaging 2010;28:1113–9*, 2010.

[94] J. T. Lizier, J. Heinzle, A. Horstmann, J. D. Haynes, and M. Prokopenko, "Multivariate information-theoretic measures reveal directed information structure and task relevant changes in fmri connectivity," *J Comput Neurosci 2011;30:85–107*, 2011.

[95] R. Cofre and C. Maldonado, "Information entropy production of maximum entropy markov chains from spike trains," *Entropy 2018, 20, 34*, 2018.

[96] N. A. Cayco-Gajic, J. Zylberberg, and E. Shea-Brown, "A moment-based maximum entropy model for fitting higher-order interactions in neural data," *Entropy 2018, 20, 489*, 2018.

[97] J. Pillow, J. Shlens, L. Paninski, A. Sher, A. M. Litke, E. J. Chichilnisky, and E. P. Simoncelli, "Spatio-temporal correlations and visual signaling in a complete neuronal population," *Nature 2008, 454, 995–999*, 2008.

[98] D. Ostwald and A. P. Bagshaw, "Information theoretic approaches to functional neuroimaging," *Magnetic Resonance Imaging, Volume 29, Issue 10, 2011,Pages 1417-1428, ISSN 0730-725X*, 2011.

[99] E. Jaynes, "Information theory and statistical mechanics," *Phys. Rev. 1957, 106, 620*, 1957.

[100] S. Si, B. Wang, X. Liu, C. Yu, C. Ding, and H. Zhao, "Brain network modeling based on mutual information and graph theory for predicting the connection mechanism in the progression of alzheimer's disease," *Entropy (Basel, Switzerland) vol. 21,3 300*, 2019.

[101] V. Griffith and C. Koch, "Quantifying synergistic mutual information. in guided self-organization: Inception," *Springer: Berlin, Germany, 2014; pp. 159–190*, 2014.

[102] W. Z. et al., "Mutual information better quantifies brain network architecture in children with epilepsy," *Computational and Mathematical Methods in Medicine (2018): n. pag*, 2018.

[103] M. Liu, X. Liu, A. Hildebrandt, and C. Zhou, "Individual cortical entropy profile: Test–retest reliability, predictive power for cognitive ability, and neuroanatomical foundation," *Cerebral Cortex Communications, Volume 1, Issue 1, 2020, tgaa015*, 2020.

[104] D. B. de Araujo, W. Tedeschi, A. C. Santos, J. Elias, U. P. C. Neves, and O. Baffa, "Shannon entropy applied to the analysis of event-related fmri time series," *Neuroimage 2003;20:311–7*, 2003.

[105] H. G. Eyherabide and I. Samengo, "Assessing the relevance of specific response features in the neural code," *Entropy 2018, 20, 879*, 2018.

[106] J. Gao, J. Hu, F. Liu, and Y. Cao, "Multiscale entropy analysis of biological signals: A fundamental bi-scaling law," *Front. Comput. Neurosci. 2015, 9, 64*, 2015.

[107] S. Keshmiri, H. Sumioka, M. Okubo, and H. Ishiguro, "An information-theoretic approach to quantitative analysis of the correspondence between skin blood flow and functional near-infrared spectroscopy measurement in prefrontal cortex activity," *Front. Neurosci. 2018, 13, 79*, 2018.

[108] G. Markowsky, "Information theory," *Encyclopedia Britannica*, 2017.

[109] W. Weaver, "Recent contributions to the mathematical theory of communication," *ETC: A Review of General Semantics, vol. 10, no. 4, 1953, pp. 261–281*, 1953.

[110] P. Perrot, "A to z of thermodynamics," *Oxford University Press. ISBN 0-19-856552-6*, 1998.

[111] Y. A. Cengel, "On entropy, information, and conservation of information," *Entropy (Basel). 2021;23(6):779. Published 2021 Jun 19. doi:10.3390/e23060779*, 2021.

[112] F. Bavaud, J. C. Chappelier, and J. Kohlas, "An introduction to information theory and applications," *Book, version 2.04 - 20050309 - UniFr course*, 2005.

[113] C. M. Bishop, "Pattern recognition and machine learning," *Springer, New York, ISBN: 978-1-4939-3843-8*, 2006.

[114] R. Frigg and C. Werndl, "Entropy - a guide for the perplexed," *Probabilities in Physics. Oxford University Press. pp. 115-142*, 2011.

[115] A. V. et al., "Shannon entropy of brain functional complex networks under the influence of the psychedelic ayahuasca," *Scientific Reports volume 7, Article number: 7388*, 2017.

[116] A. R. McIntosh, N. Kovacevic, and R. J. Itier, "Increased brain signal variability accompanies lower behavioral variability in development," *PLoS Comput. Biol. 2008, 4, e1000106*, 2008.

[117] G. Tononi, O. Sporns, and G. M. Edelman, "A measure for brain complexity: Relating functional segregation and integration in the nervous system," *Proc. Natl. Acad. Sci. USA Vol. 91, pp. 5033-5037, May 1994 Neurobiology*, 1994.

[118] M. Prokopenko and C. Gershenson, "Entropy methods in guided self-organisation," *Entropy 2014, 16, 5232-5241*, 2014.

[119] F. Cieri, X. Zhuang, J. Z. K. Caldwell, and D. Cordes, "Brain entropy during aging through a free energy principle approach," *Front Hum Neurosci. 2021;15:647513 doi:10.3389/fnhum.2021.647513*, 2021.

[120] G. Buzsaki and K. Mizuseki, "The log-dynamic brain: how skewed distributions affect network operations," *Nat Rev Neurosci. 2014 April ; 15(4): 264–278*, 2014.

[121] J. Hohwy, "The predictive mind," *Oxford University PressISBN: 978-0199686735*, 2013.

[122] K. Friston and G. Buzsaki, "The functional anatomy of time: What and when in the brain," *Trends Cogn Sci. 2016 Jul; 20(7):500-511*, 2016.

[123] A. Clark, "Busting out: Predictive brains, embodied minds, and the puzzle of the evidentiary veil," *Wiley, eBook, https://doi.org/10.1111/nous.12140*, 2016.

[124] K. Nave and A. C. et al., "Wilding the predictive brain," *WIREs Cognitive Science 2020; 11:e1542*, 2020.

[125] N. M. Timme and C.Lapish, "A tutorial for information theory in neuroscience," *eNeuro 29 June 2018, 5 (3) ENEURO 0052-18*, 2018.

[126] G. Tononi and G. M. E. et al., "Functional clustering: Identifying strongly interactive brain regions in neuroimaging data," *Neuroimage 1998, 7, 133–149*, 1998.

[127] T. O. Sharpee, A. J. Calhoun, and S. H. Chalasani, "Information theory of adaptation in neurons, behavior, and mood," *Curr. Opin. Neurobiol. 2014, 25, 47–53*, 2014.

[128] R. Brasselet and A. Arleo, "Category structure and categorical perception jointly explained by similarity-based information theory," *Entropy 2018, 20, 527*, 2018.

[129] E. Bonmati, A. Bardera, M. Feixas, and I. Boada, "Novel brain complexity measures based on information theory," *Entropy 2018, 20, 491*, 2018.

[130] A. D. et al., "Human consciousness is supported by dynamic complex patterns of brain signal coordination," *Sci. Adv. 2019, 5, eaat7603*, 2019.

[131] Z. H. et al., "Temporal circuit of macroscale dynamic brain activity supports human consciousness," *Sci. Adv. 2020, 6, eaaz0087*, 202.

[132] Y. Fan, L. L. Zeng, H. Shen, J. Qin, F. Li, and D. Hu, "Lifespan development of the human brain revealed by large-scale network eigen-entropy," *Entropy 2017, 19, 471*, 2017.

[133] Y. Yao, W. L. Lu, B. Xu, C. B. Li, C. P. Lin, D. Waxman, and J. F. Feng, "The increase of the functional entropy of the human brain with age," *Scientific Reports volume 3, Article number: 2853*, 2013.

[134] A. C. Yang, C. C. Huang, H. L. Yeh, M. E. Liu, C. J. Hong, and P. C. T. et al., "Complexity of spontaneous bold activity in default mode network is correlated with cognitive function in normal male elderly: a multiscale entropy analysis," *Neurobiol Aging. 2013 Feb;34(2):428-38. PMID: 22683008*, 2013.

[135] P. M. et al., "Analysis of fmri time-series by entropy measures," *Neuro endocrinology letters 33 5 (2012): 471-6*, 2012.

[136] W. L. Zheng and B. L. Lu, "Investigating critical frequency bands and channels for eeg-based emotion recognition with deep neural networks," *IEEE Trans. Auton. Ment. Dev. 2015, 7, 162–175*, 2015.

[137] T. Wu, K. P. Schulz, and J. Fan, "Activation of the cognitive control network associated with information uncertainty," *Neuroimage. 2021 Apr 15; 230:117703*, 2021.

[138] T. W. et al., "Supramodal mechanisms of the cognitive control network in uncertainty processing," *Cereb Cortex. 2020 Nov 3; 30(12):6336-6349*, 2020.

[139] F. Mushtaq, A. Bland, and A. Schaefer, "Uncertainty and cognitive control," *Frontiers in Psychology, Vol 2*, 2011.

[140] W. S. Pritchard, P. J. Laurienti, J. H. Burdette, and S. Hayasaka, "Functional brain networks formed using cross-sample entropy are scale free," *Brain Connect. 4(6):454-64. doi: 10.1089/brain.2013.0217. PMID: 24946057; PMCID: PMC4120805*, 2014.

[141] P. Kant, J. Hazarika, and S. H. Laskar, "Wavelet transform based approach for eeg feature selection of motor imagery data for braincomputer interfaces," *Third International Conference on Inventive Systems and Control (ICISC), 2019, pp. 101-105*, 2019.

[142] W. E. Forum, "New vision for education unlocking the potential of technology," *The Jobs Reset Summit, World Economic Forum, Geneva*, 2015.

[143] W. E. Forum, "Top 10 job skills of tomorrow – and how long it takes to learn them," *The Jobs Reset Summit*, 2020.

[144] P. A. Frensch and J. Funke, "Complex problem solving: The european perspective," *Psychology Press. https://doi.org/10.4324/9781315806723*, 1995.

[145] A. C. Ruocco, A. H. Rodrigo, J. Lam, S. I. D. D. SI, B. Graves, and H. Ayaz, "A problem-solving task specialized for functional neuroimaging: validation of the scarborough adaptation of the tower of london (s-tol) using near-infrared spectroscopy," *Front Hum Neurosci. 2014 Mar 28;8:185. doi: 10.3389/fnhum.2014.00185. PMID: 24734017; PMCID: PMC3975118*, 2014.

[146] S. D. Newman, J. A. Greco, and D. Lee, "An fmri study of the tower of london: a look at problem structure differences," *Brain Res. 2009 Aug 25;1286:123-32*, 2009.

[147] S. D. Newman and S. R. Green, "Complex problem solving," *Brain Mapping: An Encyclopedic Reference, vol. 3, pp. 543-549 Elsevier*, 2015.

[148] A. Alchihabi, O. Ekmekci, B. B. Kivilcim, S. D. Newman, and F. Y. Vural, "Analyzing complex problem solving by dynamic brain networks," *Front. Neuroinform. 15:670052 doi:10.3389/fninf.2021.670052*, 2021.

[149] D. H. Jonassen and P. Henning, "Mental models: Knowledge in the head and knowledge in the world," *Educational Technology, 39(3), 37–42*, 1999.

[150] S. D. Newman and P. C. et al., "Frontal and parietal participation in problem solving in the tower of london: fmri and computational modeling of planning and high-level perception," *Neuropsychologia, 41, 1668-1682*, 2003.

[151] E. Alpaydın, "Lecture notes for introduction to machine learning," *The MIT Press (V1.1)*, 2004.

[152] E. Parzen, "On estimation of a probability density function and mode," *The Annals of Mathematical Statistics, 33(3), 1065-1076.*, 1962.

[153] M. Rosenblatt, "Remarks on some nonparametric estimates of a density function," *The Annals of Mathematical Statistics. 27 (3): 832–837*, 1956.

[154] M. C. Jones, J. S. Marron, and S. J. Sheather, "A brief survey of bandwidth selection for density estimation," *Amer. Stat. Assoc., 91(433):401–407*, 1996.

[155] B. W. Silverman, "Density estimation for statistics and data analysis," *London: Chapman & Hall/CRC. p. 45. ISBN 978-0-412-24620-3*, 1986.

[156] S. Kullback and R. A. Leibler, "On information and sufficiency," *Annals of Mathematical Statistics. 22 (1): 79–86*, 1951.

[157] A. Alchihabi, B. B. Kivilicim, O. Ekmekci, S. D. Newman, and F. T. Y. Vural, "Decoding cognitive subtasks of complex problem solving using fmri signals," *26th Signal Processing and Communications Applications Conference (SIU) (2018): 1-4*, 2018.

[158] N. Tzourio-Mazoyer, B. Landeau, D. P. ad F. Crivello, O. Etard, N. Delcroix, B. Mazoyer, and M. Joliot, "Automated anatomical labeling of activations in spm using a macroscopic anatomical parcellation of the mni mri single-subject brain," *Neuroimage. 2002 Jan;15(1):273-89. doi: 10.1006/nimg.2001.0978. PMID: 11771995*, 2002.

[159] G. G. Degirmendereli and F. T. Y. Vural, "Estimating static and dynamic brain networks by kullback-leibler divergence from fmri data," *2020 25th International Conference on Pattern Recognition (ICPR) Milan, Italy*, 2021.

[160] G. G. Degirmendereli, S. D. Newman, and F. T. Y. Vural, "On the entropy of brain anatomic regions for complex problem solving," *IEEE 19th International Conference on Bioinformatics and Bioengineering (BIBE), 2019, pp. 603-608*, 2019.

[161] R. Lazeron, S. Rombouts, W. Machielsen, P. Scheltens, M. P. Witter, H. Uylings, and F. Barkhof, "Visualizing brain activation during planning: The tower of london test adapted for functional mr imaging," *AJNR Am J Neuroradiol 21:1407–1414*, 2000.

[162] J. M. Unterrainer, B. Rahm, C. P. Kaller, C. C. Ruff, J. Spreer, B. J. Krause, R. Schwarzwald, H. Hautzel, and U. Halsband, "When planning fails: individual differences and error-related brain activity in problem solving," *Cerebral Cortex V 14 N 12, Oxford University Press*, 2004.

[163] M. Freton, C. Lemogne, L. Bergouignan, P. Delaveau, S. Lehericy, and P. Fossati, "The eye of the self: Precuneus volume and visual perspective during autobiographical memory retrieval," *Brain structure & function. 219. 10.1007/s00429-013-0546-2*, 2013.

[164] T. Ogiso, K. Kobayashi, and M. Sugishita, "The precuneus in motor imagery," *Neuroreport. 11. 1345-9. 10.1097/00001756-200004270-00039*, 2000.

[165] J. M. Unterrainer, C. C. Ruff, B. Rahm, C. P. Kaller, J. Spreer, R. Schwarzwald, and U. Halsband, "The influence of sex differences and individual task performance on brain activation during planning," *NeuroImage 24 (2005) 586–590*, 2005.

[166] J. Fincham and J. A. et al., "Neural mechanisms of planning: A computational analysis using event-related fmri," *Proceedings of the National Academy of Science, 2002 Mar 5;99(5):3346-51*, 2002.

[167] A. Alchihabi and F. T. Y. V. et al., "A dynamic network representation of fmri for modeling and analyzing the problem solving task," *2018 IEEE 15th International Symposiumon Biomedical Imaging (ISBI 2018)*, 2018.

[168] M. Xia, J. Wang, and Y. He, "Brainnet viewer: a network visualization tool for human brain connectomics," *PloS one, vol. 8, no.7, pp. e68910*, 2013.

[169] F. Cazalis, R. Valabregue, M. Pelegrini-Issac, S. Asloun, T. W. Robbins, and S. Granon, "Individual differences in prefrontal cortical activation on the tower of london planning task: implication for effortful processing," *European Journal of Neuroscience, Vol. 17, pp. 2219–2225*, 2003.

[170] S. Newman, B. Pruce, A. Rusia, and T. Burns, "The effect of strategy on problem solving: An fmri study," *The Journal of Problem Solving: Vol. 3 : Iss. 1 , Article 2*, 2010.

[171] P. Saariluoma, "Apperception and restructuring in chess players' problem solving," *Lines of thinking (Vol. 2). New York: Wiley*, 1990.

[172] W. G. Chase and H. A. Simon, "Perception in chess," *Cognitive Psychology, 4, 55-81*, 1973.

[173] F. Gobet and H. A. Simon, "Templates in chess memory: A mechanism for recalling several boards," *Cognitive Psychology, 31, 1-40*, 1996.

[174] F. Gobet, "A pattern-recognition theory of search in expert problem solving," *Thinking and Reasoning, 3, 291-313.*, 1997.

[175] G. Polya, "Mathematics and plausible reasoning," *Princeton, NJ: Princeton University Press*, 1954.

[176] J. Larkin, J. McDermott, D. P. Simon, and H. A. Simon, "Models of competence in solving physics problems," *Cognitive Science, 4, 317-345*, 1980.

[177] M. Chi, P. Feltovich, and R. Glaser, "Categorization and representation of physics problems by experts and novices," *Cognitive Science, 5, 121-153*, 1981.

[178] A. G. Priest and R. O. Lindsay, "New light on novice-expert differences in physics problem solving," *British Journal of Psychology, 83(3), 389–405*, 1992.

[179] D. P. Simon and H. A. Simon, "Individual differences in solving physics problems," *Children's thinking: What develops? (pp. 325–348)*, 1978.

[180] L. E. Hardin, "Problem-solving concepts and theories," *Journal of veterinary medical education, 30(3), 226-229*, 2003.

[181] A. Dagher, A. M. Owen, H. Boecker, and D. J. Brooks, "Mapping the network for planning: a correlational pet activation study with the tower of london task," *Brain 1999 Oct;122 (Pt 10):1973-87.*, 1999.

[182] S. C. Baker, R. D. Rogers, A. M. Owen, C. D. Frith, R. J. Dolan, R. S. Frackowiak, and T. W. Robbins, "Neural systems engaged by planning: a pet study of the tower of london task," *Neuropsychologia, 34, 515–526*, 1996.

[183] M. T. H. Chi, , P. J. Feltovich, and R. Glaser, "Categorization and representation of physics problems by experts and novices," *Cognitive Science, Volume 5, Issue 2 April 1981 Pages 121-152*, 1981.

[184] R. L. Carhart-Harris, "The entropic brain - revisited," *Neuropharmacology, https://doi.org/10.1016/j.neuropharm.2018.03.010*, 2018.

[185] P. Duan, "Information theory-based approaches for causality analysis with industrial applications," *Ph.D. thesis, Department of Electrical & Computer Engineering, University of Alberta*, 2014.

[186] T. Schreiber, "Measuring information transfer," *Physical Review Letters, 85(2):461–464*, 2000.

[187] M. P. et al., "Synchronization as adjustment of information rates: detection from bivariate time series," *Phys. Rev. E 63 (2001) 046211*, 2001.

[188] H. H. et al., "Causal visual interactions as revealed by an information theoretic measure and fmri," *NeuroImage 31 (2006) 1051–1060*, 2006.

[189] M. Pflieger, "Using conditional mutual information to approximate causality for multivariate physiological time series," *International Journal of Bioelectromagnetism. 7.*, 2005.

[190] R. V. et al., "Transfer entropy-a model-free measure of effective connectivity for the neurosciences," *Journal of Computational Neuroscience, 30(1):45–67, 2011*, 2011.

[191] S. Palachy, "Inferring causality in time series data - a concise review of the major approaches," *Towards Data Science (towardsdatascience.com)*, 2019.

[192] C. J. Q. et al., "Estimating the directed information to infer causal relationships in ensemble neural spike train recordings," *Journal of Computational Neuroscience, 30(1):17-44. doi: 10.1007/s10827-010-0247-2*, 2011.

[193] J. W. et al., "Causal entropies—a measure for determining changes in the temporal organization of neural systems," *Journal of Neuroscience Methods, Volume 162, Issues 1–2,Pages 320-332,ISSN 0165-270*, 2007.

[194] P. O. Amblard and O. J. J. Michel, "On directed information theory and granger causality graphs," *Journal of Computational Neuroscience 30, 7–16*, 2011.

[195] I. V. I and D. Kugiumtzis, "Nonuniform state-space reconstruction and coupling detection," *Physical Review E Statistical Nonlinear & Soft Matter Physics. 2010; 82(1 Pt 2): 016207*, 2010.

[196] D. Kugiumtzis, "Direct coupling information measure from non-uniform embedding," *Physical Review E. 2013; 87(6): 062918*, 2013.

[197] X. W. X and L. Xu, "A study for multiscale information transfer measures based on conditional mutual information," *PLoS ONE 13(12): e0208423*, 2018.

[198] Y. O. et al, "A lognormal recurrent network model for burst generation during hippocampal sharp waves," *Journal of Neuroscience 28 October 2015, 35 (43) 14585-14601*, 2015.

[199] G. Scheler, "Logarithmic distributions prove that intrinsic learning is hebbian," *F1000Research 2017, 6:1222*, 2017.

[200] E. Limpert, W. A. Stahel, and M. Abbt, "Log-normal distributions across the sciences: Keys and clues," *BioScience, Volume 51, Issue 5, May 2001, Pages 341-352*, 2001.

[201] S. M. Silverstein, M. Wibral, and W. A. Phillips, "Implications of information theory for computational modeling of schizophrenia," *Computational Psychiatry. 2017; 1:82–101*, 2017.

[202] A. E. L. et al., "Utilizing mutual information analysis to explore the relationship between gray and white matter structural pathologies in schizophrenia," *Schizophr. Bull. 2018 doi: 10.1093/schbul/sby028*, 2018.

[203] L. Loued-Khenissi and K. Preuschoff, "Information theoretic characterization of uncertainty distinguishes surprise from accuracy signals in the brain," *Frontiers in Artificial Intelligence 3:5. doi: 10.3389/frai.2020.00005*, 2020.

[204] K. Friston, "The history of the future of the bayesian brain," *Neuroimage. 2012;62(2):1230-1233. doi:10.1016/j.neuroimage.2011.10.004*, 2012.

[205] J. Fan, "An information theory account of cognitive control," *Front Hum Neurosci. 2014;8 680. doi:10.3389/fnhum.2014.00680. PMID: 25228875; PMCID: PMC4151034*, 2014.

CURRICULUM VITAE

PERSONAL INFORMATION

Surname, Name: Degirmendereli Gunal, Gonul

EDUCATION

2013-2021	METU Informatics Institute, PhD Candidate in Cognitive Science
2003–2005	Bogazici University, Engineering Faculty, MS in Engineering and Technology Management
1978–1983	Middle East Technical University, Engineering Faculty, BS in Computer Engineering,

PROFESSIONAL EXPERIENCE

- Freelance SAS Consultant, Data Quality, BI (2007–2013)
- Principal Consultant & Project Manager (SAS Turkey, 2005–2007)
- Freelance Consultant, DW, BI, MIS (2002–2005)
- Project Leader, Data Warehouse (DW), BI (I-BIMSA, 2000–2001)
- Senior Consultant, Business Intelligence (BI) (I-BIMSA, 1997–1999)
- Software Marketing Support Specialist (IBM Contract, 1995–1996)
- Senior Database & Appl. Supp. Specialist (I-BIMSA, 1989–1995)
- System Analyst & Application Programmer (I-BIMSA, 1986–1989)
- Education Specialist (IBM Turkey, 1983–1985)

PUBLICATIONS

- G. Gunal Degirmendereli, F. T. Yarman Vural, "Representation of Cognitive States by Shannon Entropy," 2021 29th Signal Processing and

Communications Applications Conference (SIU), 2021, pp. 1-4, doi: 10.1109/SIU53274.2021.9477943.

- G. Gunal Degirmendereli, F.T.Yarman Vural, "Estimating Static and Dynamic Brain Networks by Kullback-Leibler Divergence from fMRI Data," 2020 25th International Conference on Pattern Recognition (ICPR) Milan, Italy, 2021, pp. 5913-5919, doi: 10.1109/ICPR48806.2021.9413047.
- G. Gunal Degirmendereli, S. D. Newman and F. T. Yarman Vural, "On the Entropy of Brain Anatomic Regions for Complex Problem Solving," IEEE 19th International Conference on Bioinformatics and Bioengineering (BIBE), 2019, pp. 603-608, doi: 10.1109/BIBE.2019.00115.
- G. Degirmendereli, A. O. Erdem, A new Approach to Dynamic Plan Selection, DB2 Update Journal, XEPHON PLC, May 1992

PAPER PRESENTATIONS

- G.G Degirmendereli, F.Y. Vural, "Representation of Cognitive States by Shannon Entropy - Bilişsel Durumların Shannon Entropisi ile Temsili" 29th Signal Processing and Communications Applications Conference - 29. IEEE Sinyal Isleme ve Iletisim Uygulamalari Kurultayi (SIU 2021), Bahcesehir University, Istanbul, Turkey
- G.G. Degirmendereli, F.Y. Vural, "Karmasik problem cözme sürecinde beyin anatomik bölgelerinin entropi analizi - On the entropy of brain anatomic regions for complex problem solving", 18th National Neuroscience Congress - 18. Ulusal Sinirbilim Kongresi (USK 2020), Bilkent University, Ankara, Turkey
- G.G. Degirmendereli, F.Y. Vural, "Estimating Static and Dynamic Brain Networks by Kullback-Leibler Divergence from fMRI Data", 25th International Conference on Pattern Recognition (ICPR 2020) Milan, Italy.
- G.G. Degirmendereli, S.D. Newman and F.Y. Vural, “On the Entropy of Brain Anatomic Regions for Complex Problem Solving”, 7th International Symposium on Brain and Cognitive Science (ISBCS 2020), Bilkent University, Ankara, Turkey.
- G.G. Degirmendereli, S.D. Newman and F.Y. Vural, "On the Entropy of Brain Anatomic Regions for Complex Problem Solving," IEEE 19th International Conf. on Bioinformatics and Bioengineering (BIBE 2019), Athens, Greece.
- G. Degirmendereli, N. Demirsoy, “The Story of Building a Data Warehouse”, Bilisim’2000 Informatics Congress, Istanbul, Turkey
- G. Degirmendereli, “Building a Data Warehouse”, Bilisim’98 Informatics Congress, Istanbul, Turkey

- G. Degirmendereli, “Data Warehouse Solution”, Open System’98 Symposium, 1998, Istanbul, Turkey
- N. Akmeric, G. Degirmendereli, “The Place and Importance of Information in Contemporary Business Life”, 7th Turkey Computer Congress, 1993, Istanbul, Turkey
- Bingol O., Degirmendereli G., Ergin D., “What Do We Expect from The Education of Computer Professionals”, 4th Turkey Computer Congress 1990, Istanbul, Turkey

AWARDS AND FELLOWSHIPS

- International Association for Pattern Recognition (IAPR) "Best Student Paper Award" in Image and Signal Processing track at the “25th International Conference on Pattern Recognition - ICPR 2020” Milano, Italy, January 2021
- Best Paper Award, 4th Turkey Computer Congress 1990
- Honor Award, Scientific and Technological Research Council of Turkey (TÜBİTAK), Middle School Mathematics Competition 1974
- Scientific and Technological Research Council of Turkey (TÜBİTAK) undergraduate scolarship (1977-1980)

PROFESSIONAL COURSES GIVEN

- Dataflux dfPower Customize Training (2010 Turkcell, 2007 Halkbank)
- Dataflux dfPower Studio Training (2008 SSI; 2008 Akbank; 2009 Turkcell; 2010 Kuveyt Turk Bank)
- Introduction to SAS Business Intelligence Applications (2006, TUİK)
- Creating and Exploiting OLAP Using the SAS System (2005 AVIVASA; 2006 TUİK)
- Data Warehouse Enablement–Techniques & Methods. (2003 Is Bank)
- DB2 UDB Administration Workshop for Unix (2002, Erdemir)
- DB2 DBA Workshop for Common Server (1996-1997, IBM Istanbul)
- Data Replication (Data Propagator Relational for UNIX) (1996, IBM Istanbul)
- DB2 Database Design (1996, IBM Istanbul)
- DB2 for MVS Application Design (1995, IBM Istanbul)
- DB2 Application Programming (1993, I-BIMSA Istanbul)
- The SQL Workshop (1993, I-BIMSA Istanbul)

TEZ İZİN FORMU / THESIS PERMISSION FORM

ENSTİTÜ / INSTITUTE

Fen Bilimleri Enstitüsü / Graduate School of Natural and Applied Sciences [ ]

Sosyal Bilimler Enstitüsü / Graduate School of Social Sciences [ ]

Uygulamalı Matematik Enstitüsü / Graduate School of Applied Mathematics [ ]

Enformatik Enstitüsü / Graduate School of Informatics [X]

Deniz Bilimleri Enstitüsü / Graduate School of Marine Sciences [ ]

YAZARIN / AUTHOR

Soyadı / Surname : GUNAL DEGIRMENDERELI
Adı / Name : GONUL
Bölümü / Department : COGNITIVE SCIENCE

TEZİN ADI / TITLE OF THE THESIS (**İngilizce** / English) : AN INFORMATION THEORETIC REPRESENTATION OF HUMAN BRAIN FOR DECODING MENTAL STATES OF COMPLEX PROBLEM SOLVING

TEZİN TÜRÜ / DEGREE: **Yüksek Lisans** / Master [ ] **Doktora** / PhD [X]

1. **Tezin tamamı dünya çapında erişime açılacaktır. /** Release the entire work immediately for access worldwide. [X]
2. **Tez iki yıl süreyle erişime kapalı olacaktır.** / Secure the entire work for patent and/or proprietary purposes for a period of **two year. *** [ ]
3. **Tez altı ay süreyle erişime kapalı olacaktır.** / Secure the entire work for period of **six months**. ***** [ ]

***** *Enstitü Yönetim Kurulu Kararının basılı kopyası tezle birlikte kütüphaneye teslim edilecektir. A copy of the Decision of the Institute Administrative Committee will be delivered to the library together with the printed thesis.*

Yazarın imzası / Signature **Tarih** / Date