

Contents

| | |
|---|----------|
| <i>List of Figures</i> | page xii |
| <i>List of Tables</i> | xv |
| <i>Preface</i> | xvii |
| <i>Acknowledgments</i> | xix |
| PART I. INTRODUCTION | |
| 1 Sequence Analysis in the Social Sciences | 3 |
| 1.1 Timing and Context of the Book | 4 |
| 1.2 Contributions | 7 |
| 1.3 Audience and Scope | 9 |
| 1.4 Related References | 14 |
| 1.5 Plan of the Book | 16 |
| PART II. THEORETICAL BACKGROUND | |
| 2 Theoretical Foundations of Social Sequence Analysis | 21 |
| 2.1 What Are Social Sequences? | 21 |
| 2.2 The Structural Origins of Social Sequences | 22 |
| 2.3 Why Social Sequences Matter | 25 |
| 2.4 Origins of Social Sequence Analysis Techniques | 32 |
| 2.4.1 <i>Narrative Positivism</i> | 33 |
| 2.4.2 <i>The Relational Nature of Social Phenomena</i> | 36 |
| 2.4.3 <i>The Development of Whole-Sequence Analysis Methods</i> | 38 |
| 2.4.4 <i>Pivotal Criticisms</i> | 41 |
| 2.5 The Second Wave of Social Sequence Analysis | 45 |
| 2.5.1 <i>Sequences as Networks</i> | 46 |
| 2.5.2 <i>Microsequence Analysis</i> | 52 |
| 2.6 Looking Ahead | 55 |

PART III. SOCIAL SEQUENCE ANALYSIS CONCEPTS
AND TECHNIQUES

| | | |
|-------|---|----|
| 3 | Sequence Analysis Concepts and Data | 59 |
| 3.1 | Sequence Properties | 59 |
| 3.1.1 | <i>Positions</i> | 60 |
| 3.1.2 | <i>Elements</i> | 60 |
| 3.1.3 | <i>Substructures</i> | 61 |
| 3.1.4 | <i>Subjects</i> | 62 |
| 3.2 | Key Issues in Sequence Data | 62 |
| 3.2.1 | <i>Recurrence</i> | 63 |
| 3.2.2 | <i>Ties and Multidimensionality</i> | 65 |
| 3.2.3 | <i>Gaps</i> | 66 |
| 3.3 | The Sequence Universe | 67 |
| 3.3.1 | <i>Size</i> | 67 |
| 3.3.2 | <i>Boundary Specification</i> | 69 |
| 3.4 | Sequence Data | 70 |
| 3.4.1 | <i>Data Collection</i> | 70 |
| 3.4.2 | <i>Element-Position Sampling</i> | 73 |
| 3.4.3 | <i>Units of Measure</i> | 74 |
| 3.5 | Data Used in This Book | 75 |
| 3.5.1 | <i>The Survey of Health, Ageing, and Retirement in Europe (SHARE)</i> | 75 |
| 3.5.2 | <i>The Correlates of War Global Trade Network</i> | 76 |
| 3.5.3 | <i>The Davis, Gardner, and Gardner Deep South Study</i> | 78 |
| 3.5.4 | <i>The Multinational Time Use Study (MTUS)</i> | 79 |
| 3.5.5 | <i>The American Time Use Survey (ATUS)</i> | 80 |
| 4 | Detecting Sequence Structure | 83 |
| 4.1 | Descriptive Analysis | 83 |
| 4.1.1 | <i>Sequence Length</i> | 84 |
| 4.1.2 | <i>Element Frequencies</i> | 84 |
| 4.1.3 | <i>n-grams</i> | 85 |
| 4.1.4 | <i>Position Reports</i> | 85 |
| 4.2 | Describing Stochastic Patterns | 86 |
| 4.2.1 | <i>Transition Matrices</i> | 86 |
| 4.2.2 | <i>Markov Chains</i> | 91 |
| 4.3 | Sequential Connection | 91 |
| 4.3.1 | <i>Empirical Illustration: Gender Differences in the Parenthood-Stress Link</i> | 93 |
| 4.4 | Stationarity | 94 |
| 4.4.1 | <i>Empirical Illustration: Stationarity in the World System</i> | 96 |
| 4.5 | Spells | 97 |

| | | |
|-----|---|-----|
| 4.6 | Homogeneity | 98 |
| 4.7 | On Using Summary Statistics and Tests | 99 |
| 4.8 | Visual Aids | 100 |
| | 4.8.1 <i>Transition Bubble Graphs</i> | 100 |
| | 4.8.2 <i>State Transition Diagrams</i> | 101 |
| | 4.8.3 <i>Sequence Index Plots</i> | 102 |
| | 4.8.4 <i>State Distribution Graphs</i> | 104 |
| | 4.8.5 <i>Tempograms</i> | 104 |
| | 4.8.6 <i>Sequence-Network Diagrams</i> | 105 |
| | 4.8.7 <i>On the Use of Color and Grayscale</i> | 105 |
| 4.9 | Looking Ahead | 108 |
| 5 | Whole-Sequence Comparison Methods | 109 |
| 5.1 | Sequence Alignment | 110 |
| | 5.1.1 <i>Sequence Alignment Operations</i> | 110 |
| | 5.1.2 <i>Operation Costs and Distance</i> | 111 |
| 5.2 | Classical Optimal Matching (OM) | 111 |
| | 5.2.1 <i>Finding the Optimal Solution</i> | 113 |
| | 5.2.2 <i>The Needleman-Wunsch Algorithm</i> | 114 |
| 5.3 | Basic Operation Cost Regimes | 115 |
| | 5.3.1 <i>Levenshtein Distance</i> | 115 |
| | 5.3.2 <i>Levenshtein II Distance</i> | 115 |
| | 5.3.3 <i>Hamming Distance</i> | 116 |
| 5.4 | Considerations in Setting Operation Costs | 116 |
| | 5.4.1 <i>Order versus Timing</i> | 116 |
| | 5.4.2 <i>Time Warping</i> | 118 |
| 5.5 | Criticisms of Classical OM | 119 |
| | 5.5.1 <i>Theory/Method Fit</i> | 120 |
| | 5.5.2 <i>Arbitrary Operation Costs</i> | 120 |
| 5.6 | Improvements on Classical OM | 122 |
| | 5.6.1 <i>Variable Substitution Costs</i> | 122 |
| | 5.6.2 <i>Distance Normalization</i> | 124 |
| | 5.6.3 <i>Reference Sequence Comparison</i> | 125 |
| | 5.6.4 <i>Spell-Adjusted Distances</i> | 126 |
| | 5.6.5 <i>The Dynamic Hamming Method</i> | 127 |
| 5.7 | Detecting Whole Sequence Patterns | 129 |
| | 5.7.1 <i>Hierarchical Cluster Analysis</i> | 130 |
| | 5.7.2 <i>Identifying Sequence Classes</i> | 133 |
| | 5.7.3 <i>Assessing Validity and Reliability</i> | 136 |
| | 5.7.4 <i>Describing Sequence Classes</i> | 139 |
| 5.8 | Recent Advances in Sequence Comparison | 143 |
| | 5.8.1 <i>Multidimensional Sequence Analysis</i> | 144 |
| | 5.8.2 <i>Two-Stage Optimal Matching (2SOM) Analysis</i> | 146 |
| | 5.8.3 <i>Transition Sequence Analysis</i> | 147 |

| | | |
|---|---|-----|
| 5.8.4 | <i>Nonalignment Techniques</i> | 148 |
| 5.8.5 | <i>Beyond Cluster Analysis</i> | 150 |
| PART IV. NEW DIRECTIONS IN SOCIAL SEQUENCE ANALYSIS | | |
| 6 | Network Methods for Sequence Analysis | 155 |
| 6.1 | Theoretical Rationale | 156 |
| 6.1.1 | <i>Networks of Sequences</i> | 156 |
| 6.1.2 | <i>Sequences as Bases of Affiliations</i> | 160 |
| 6.2 | Network Concepts and Terms | 163 |
| 6.2.1 | <i>Network Matrices and Components</i> | 163 |
| 6.2.2 | <i>Directed and Undirected Networks</i> | 164 |
| 6.2.3 | <i>One- and Two-Mode Networks</i> | 165 |
| 6.2.4 | <i>Subject Comembership and Element Overlap</i> | 165 |
| 6.3 | Sequence-Network Construction | 167 |
| 6.3.1 | <i>Nonrecurrent Sequence Networks</i> | 167 |
| 6.3.2 | <i>Recurrent Sequence Affiliation Networks</i> | 170 |
| 6.3.3 | <i>Directed Sequence Networks</i> | 171 |
| 6.3.4 | <i>Multidimensional Sequence Networks</i> | 172 |
| 6.4 | Visualization | 172 |
| 6.4.1 | <i>Relationships among Subjects</i> | 173 |
| 6.4.2 | <i>Two-Mode Ordered Event Networks</i> | 174 |
| 6.4.3 | <i>Two-Mode Recurrent Sequences</i> | 176 |
| 6.4.4 | <i>Directed Sequence Networks</i> | 179 |
| 6.5 | Structural Measures for Sequence Networks | 184 |
| 6.6 | Identification of Subject/Element Subsets | 189 |
| 6.6.1 | <i>Subsets in Sequence Affiliation Networks</i> | 190 |
| 6.6.2 | <i>Element Subsets in Directed Sequence Networks</i> | 193 |
| 6.6.3 | <i>Sequence Motifs</i> | 195 |
| 6.7 | Event Structure Analysis | 196 |
| 6.8 | Statistical Network Models for Sequential Events | 200 |
| 6.8.1 | <i>Exponential Random Graph (p^*) Models</i> | 200 |
| 6.8.2 | <i>ERGMs for Ordered Nonrecurrent Sequence Networks</i> | 201 |
| 6.9 | Dynamic Sequence Networks | 202 |
| 6.9.1 | <i>Subject- and Element-Level Change</i> | 203 |
| 6.9.2 | <i>Sequence Subset Evolution</i> | 206 |
| 6.9.3 | <i>Whole Sequence-Network Evolution</i> | 207 |
| 6.10 | Conclusion | 209 |
| 7 | Social Microsequence Analysis | 210 |
| 7.1 | An Exemplary Sequence Context | 211 |
| 7.2 | The Elements of Social Microsequences | 214 |
| 7.2.1 | <i>Relationship to Microinteraction Sequences</i> | 215 |

| | | |
|---|--|-----|
| 7.2.2 | Elements and Positions | 215 |
| 7.2.3 | Data Collection and Availability | 216 |
| 7.3 | Nonnetwork Sequence Approaches | 217 |
| 7.3.1 | OM Sequence Classification | 220 |
| 7.3.2 | Transition and Switching Analysis | 222 |
| 7.4 | Sequence-Network Approaches | 226 |
| 7.4.1 | Synchrony | 227 |
| 7.4.2 | Measurement of Synchrony | 230 |
| 7.4.3 | Identifying Sources of Synchrony | 238 |
| 7.4.4 | Routine | 241 |
| 7.4.5 | Measurement and Visualization of Routine | 244 |
| 7.5 | Next Steps | 251 |
| PART V. CONCLUSIONS | | |
| 8 | The Promise of Social Sequence Analysis | 255 |
| 8.1 | Limitations | 256 |
| 8.2 | Future Research | 257 |
| 8.2.1 | Routine and Routinization | 257 |
| 8.2.2 | Sequence Networks, Network Sequences | 258 |
| 8.2.3 | Sequential Statistical Inference | 259 |
| 8.2.4 | Data Collection | 261 |
| Appendix A Recent Whole-Sequence Pattern Analyses | | 263 |
| Appendix B Linkage Criteria for Agglomerative Hierarchical Clustering | | 270 |
| References | | 275 |
| Index | | 303 |