A MULTIMODAL ANALYSIS OF SYNCHRONY DURING DYADIC INTERACTION USING A METRIC BASED ON SEQUENTIAL PATTERN MINING

Anil Jakkam and Carlos Busso
Multimodal Signal Processing (MSP) Laboratory
Erik Jonsson School of Engineering & Computer Science
University of Texas at Dallas

MOTIVATION

Background:
- Adaptation of interlocutors in human conversations is called synchrony, entrainment, or mimicry
- Synchrony has been studied well within single modality
- synchrony across modalities is an open question
- Measurement and quantification of synchrony could:
  - improve existing spoken dialogue systems
  - Improve emotion recognition systems

Proposed Solution:
- We use sequential pattern mining to study the role of synchrony in dyadic conversations
- We study synchrony at a turn level considering the acoustic and text modalities on the Fisher’s Corpus

Finding Relevant Sequences

Selection of Relevant Sequences
- SPADE algorithm (SPMF)
- Step 1: Discovers frequent sequences in training set
  - remove sequences with low support
  - contains over 1000 sequences

<table>
<thead>
<tr>
<th>Seq. #</th>
<th>Sequence</th>
<th>SUP</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>&lt;(9),(9)&gt;</td>
<td>0.166</td>
</tr>
<tr>
<td>2</td>
<td>&lt;(14),(9)&gt;</td>
<td>0.181</td>
</tr>
<tr>
<td>3</td>
<td>&lt;(9),(14)&gt;</td>
<td>0.149</td>
</tr>
<tr>
<td>4</td>
<td>&lt;(36),(9)&gt;</td>
<td>0.154</td>
</tr>
<tr>
<td>5</td>
<td>&lt;(36),(9)&gt;</td>
<td>0.138</td>
</tr>
<tr>
<td>6</td>
<td>&lt;(14,36),(9)&gt;</td>
<td>0.133</td>
</tr>
<tr>
<td>7</td>
<td>&lt;(36),(9)&gt;</td>
<td>0.126</td>
</tr>
<tr>
<td>8</td>
<td>&lt;(14),(14)&gt;</td>
<td>0.124</td>
</tr>
<tr>
<td>9</td>
<td>&lt;(9),(36)&gt;</td>
<td>0.122</td>
</tr>
<tr>
<td>10</td>
<td>&lt;(9),(14)&gt;</td>
<td>0.113</td>
</tr>
</tbody>
</table>

- Sequences may not inform about synchrony (e.g., <(9),(9)>)

Fisher’s corpus
- Dyadic interactions, with randomly assigned speakers
- 90 sessions (30 each for training validation and testing)

Sequential Pattern Mining
- Find frequent sequence of events
- Definitions:
  - event (e_i): relevant observations
  - itemset (i_i): unordered list of events (e_1, e_2,..., e_n)
  - sequence: ordered list of itemsets <1,2,...,n>
  - support: # of data sequences containing a given sequence

<table>
<thead>
<tr>
<th>Sequence</th>
<th>#</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;(9),(9)&gt;</td>
<td>0.128</td>
</tr>
<tr>
<td>&lt;(14),(9)&gt;</td>
<td>0.133</td>
</tr>
<tr>
<td>&lt;(9),(10)&gt;</td>
<td>0.138</td>
</tr>
<tr>
<td>&lt;(10),(9)&gt;</td>
<td>0.144</td>
</tr>
<tr>
<td>&lt;(36),(9)&gt;</td>
<td>0.149</td>
</tr>
<tr>
<td>&lt;(9),(14)&gt;</td>
<td>0.154</td>
</tr>
</tbody>
</table>

Example:
- Seq. <(a),(b)> has support 4

Future Work:
- Using other modalities
- Extension to multiparty interaction
- Incorporate a variable window, rather than just considering the adjacent turns

DISCUSSION

Conclusions:
- This framework captures the local interplay of multiple modalities that lead to synchrony
- Sequential pattern mining is a fast and efficient way to discover frequent sequences
- We developed a metric that effectively represents synchrony

Future Work:
- Use the sequential features to classify engagement, depression, or empathy.
- Incorporate a variable window, rather than just considering the adjacent turns
- Extension to multiparty interaction
- Using other modalities

Acknowledgment: This study was funded by NSF (grant IIS1217104) and a NSF CAREER award IIS-1453781