Dependency and Structure in Pattern Recognition

Robert M. Haralick

Computer Science, Graduate Center
City University of New York
Dependency and Structure

- Pattern Classification
  - Dependency between observations and classes
- Prediction
  - Dependency between the *independent* variables and the response variable
- Compact Representations
  - Dimensionality Reduction
  - Manifold Learning
Measuring Strength of Dependency between/among Variables

Determining the Dependency Constraint
  - Probabilistic Dependency
  - Coherence: Between Values
Kinds of Dependencies

Graphics from Jason Noble, University of Southampton
For some values of $X$, $Y$ has multiple values
  $Y$ is not a function of $X$

For some values of $Y$, $X$ has multiple values
  $Y$ is not a function of $X$

There are two non-linear disconnected manifolds
Any patch of an image that shows a texture is a region having a stochastic dependency among the pixel values of the patch.

- Gray level co-occurrence matrix
  - Distance
  - Angle

- Functionals of the co-occurrence matrix can be used as features in distinguishing one texture from another
\[ N_1 = \{ ((r, c), (u, v)) \in (R \times C)^2 \mid (u, v) = (r - 1, c + 1) \text{ or } (u, v) = (r + 1, c - 1) \} \]
\[ N_2 = \{ ((r, c), (u, v)) \in (R \times C)^2 \mid (u, v) = (r, c + 1) \text{ or } (u, v) = (r, c - 1) \} \]
\[ N_3 = \{ ((r, c), (u, v)) \in (R \times C)^2 \mid (u, v) = (r - 1, c - 1) \text{ or } (u, v) = (r + 1, c + 1) \} \]
\[ N_4 = \{ ((r, c), (u, v)) \in (R \times C)^2 \mid (u, v) = (r - 1, c) \text{ or } (u, v) = (r + 1, c) \} \]
Gray Level Cooccurrence: Major Diagonal

\[ N_1 = \{ ((r, c), (u, v)) \in (R \times C)^2 | (u, v) = (r - 1, c + 1) \text{ or } (u, v) = (r + 1, c - 1) \} \]

\[ P_1(i, j) = \frac{\#\{ ((r, c), (u, v)) \in N_1 | I(r, c) = i \text{ and } I(u, v) = j \}}{\#N_1} \]
Gray Level Cooccurrence: Left-Right

\[ N_2 = \{ ((r, c), (u, v)) \in (R \times C)^2 \mid (u, v) = (r, c + 1) \text{ or } (u, v) = (r, c - 1) \} \]

\[ P_2(i, j) = \frac{\# \{ ((r, c), (u, v)) \in N_2 \mid I(r, c) = i \text{ and } I(u, v) = j \}}{\# N_2} \]
Gray Level Cooccurrence: Minor Diagonal

\[ N_3 = \{((r, c), (u, v)) \in (R \times C)^2 \mid (u, v) = (r - 1, c - 1) \text{ or } (u, v) = (r + 1, c + 1)\} \]

\[ P_3(i, j) = \frac{\#\{((r, c), (u, v)) \in N_3 \mid I(r, c) = i \text{ and } I(u, v) = j\}}{\#N_3} \]
Gray Level Cooccurrence: Top Bottom

\[
N_4 = \{((r, c), (u, v)) \in (R \times C)^2 \mid (u, v) = (r - 1, c) \text{ or } (u, v) = (r + 1, c)\}
\]

\[
P_4(i, j) = \frac{\#\{(r, c), (u, v) \in N_4 \mid I(r, c) = i \text{ and } I(u, v) = j\}}{\#N_4}
\]
Local Neighborhoods

\[ N_1(r, c) = \{(u, v) \in R \times C \mid ((r, c), (u, v)) \in N_1\} \]

\[ N_2(r, c) = \{(u, v) \in R \times C \mid ((r, c), (u, v)) \in N_2\} \]

\[ N_3(r, c) = \{(u, v) \in R \times C \mid ((r, c), (u, v)) \in N_3\} \]

\[ N_4(r, c) = \{(u, v) \in R \times C \mid ((r, c), (u, v)) \in N_4\} \]
Correlation Feature

Because of the symmetry, $P_k(i,j) = P_k(j,i)$ and $I = J$

\[
P_{\{k,\text{row}\}}(i) = \sum_{j=1}^{J} P_k(i,j)
\]

\[
P_{\{k,\text{col}\}}(j) = \sum_{i=1}^{I} P_k(i,j)
\]

\[
\mu_k = \sum_{i=1}^{I} i P_{\{k,\text{row}\}}(i) = \sum_{j=1}^{J} j P_{\{k,\text{col}\}}(j)
\]

\[
\sigma_k^2 = \sum_{i=1}^{I} (i - \mu_k)^2 P_{\{k,\text{row}\}}(i) = \sum_{j=1}^{J} (j - \mu_k)^2 P_{\{k,\text{col}\}}(j)
\]

\[
\rho_k = \sum_{i=1}^{I} \sum_{j=1}^{J} \frac{(i - \mu_k)(j - \mu_k)}{\sigma_k^2} P_k(i,j)
\]

\[
\rho = \sum_{k=1}^{4} \rho_k \theta_k
\]
Entropy Texture Features

\[ E_{1k} = \sum_{i=1}^{I} \sum_{j=1}^{J} P_k^2(i, j) \]

\[ E_{2k} = -\sum_{i=1}^{I} \sum_{j=1}^{J} P_k(i, j) \log P_k(i, j) \]

\[ E_1 = \sum_{k=1}^{K} E_{1k} \theta_k \]

\[ E_2 = \sum_{k=1}^{K} E_{2k} \theta_k \]

Let $\alpha > 0$ and $\alpha \neq 1$, then

$$H_\alpha(p_1, \ldots, p_N) = \frac{1}{1 - \alpha} \log \left( \sum_{n=1}^{N} p_n^\alpha \right)$$

satisfies the entropy postulates. And

$$\lim_{\alpha \to 1} H_\alpha(p_1, \ldots, p_N) = - \sum_{n=1}^{N} p_n \log p_n$$

Correlation and Maximal Correlation

$$\rho(X, Y) = \frac{E[(X - \mu_X)(Y - \mu_Y)]}{\sigma_X \sigma_Y}$$

If $E[X] = E[Y] = 0$; and $V[X] = V[Y] = 1$ then
$$\rho(X, Y) = E[XY]$$

Let

$$F = \{ f : \mathbb{R} \rightarrow \mathbb{R} | E[f(X)] = 0; V[f(X)] = 1 \}$$
$$G = \{ g : \mathbb{R} \rightarrow \mathbb{R} | E[g(Y)] = 0; V[g(Y)] = 1 \}$$

Define Maximal Correlation $\rho_{max}$ by

$$\rho_{max}(X, Y) = \sup_{f \in F, g \in G} E[f(X)g(Y)]$$


Maximal Correlation Feature

The normalized joint probability matrix $Q_k = (q_k(i,j))$

$$q_k(i,j) = \frac{P_k(i,j)}{\sqrt{P_{k,\text{row}}(i)} \sqrt{P_{k,\text{col}}(j)}}$$

- The second singular value of $Q_k$: $\lambda_{k,2}$
- The maximal correlation coefficient: $\rho_{\text{max},k} = \lambda_{k,2}$

Contrast and Inverse Contrast

\[ c_k = \sum_{i=1}^{I} \sum_{j=1}^{J} |i - j|^\beta P_k(i, j) \]

\[ d_k = \sum_{i=1}^{I} \sum_{j=1}^{J} \frac{1}{1 + \alpha |i - j|} P_k(i, j) \]

\( I \) Input Image

\( P_k, k = 1, 2, 3, 4 \) Cooccurrence Probabilities

\( N_k, k = 1, 2, 3, 4 \) Local Neighborhoods

\( J \) Output Image

\[
J(r, c) = \sum_{k=1}^{4} \sum_{(u,v) \in N_k(r,c)} P_k(I(r,c), I(u,v)) \theta_k
\]

Neighborhood Joint Probability: \( P(l(u, v) : (u, v) \in N(r, c)) \)

\[
N(r, c) = \{(r, c)\} \cup \bigcup_{k=1}^{4} N_k(r, c)
\]
Neighborhood Joint Probability

\[
P(I(u, v) : (u, v) \in N(r, c)) = P(I(r, c)) \prod_{(u, v) \in N_1(r, c)} P_1(I(u, v) \mid I(r, c)) \prod_{(u, v) \in N_2(r, c)} P_2(I(u, v) \mid I(r, c)) \prod_{(u, v) \in N_3(r, c)} P_3(I(u, v) \mid I(r, c)) \prod_{(u, v) \in N_4(r, c)} P_4(I(u, v) \mid I(r, c))
\]
Neighborhood Joint Probability

\[
P(I(r, c), I(u, v) : (u, v) \in N(r, c)) = P(I(r, c)) \prod_{(u,v)\in N_1(r,c)} P_1(I(u, v) | I(r, c)) \prod_{(u,v)\in N_2(r,c)} P_2(I(u, v) | I(r, c)) \prod_{(u,v)\in N_3(r,c)} P_3(I(u, v) | I(r, c)) \prod_{(u,v)\in N_4(r,c)} P_4(I(u, v) | I(r, c))
\]

\[
J(r, c) = \log(P(I(u, v) : (u, v) \in N(r, c))) = \log(P(I(r, c))) + \sum_{(u,v)\in N_1(r,c)} \log(P_1(I(u, v) | I(r, c))) + \sum_{(u,v)\in N_2(r,c)} \log(P_2(I(u, v) | I(r, c))) + \sum_{(u,v)\in N_3(r,c)} \log(P_3(I(u, v) | I(r, c))) + \sum_{(u,v)\in N_4(r,c)} \log(P_4(I(u, v) | I(r, c)))
\]
Cooccurrence Distance 2 Relations

- $(r-2, c-2)$
- $(r-2, c-1)$
- $(r-2, c)$
- $(r-2, c+1)$
- $(r-2, c+2)$
- $(r-1, c-2)$
- $(r-1, c-1)$
- $(r-1, c)$
- $(r-1, c+1)$
- $(r-1, c+2)$
- $(r, c-2)$
- $(r, c-1)$
- $(r, c)$
- $(r, c+1)$
- $(r, c+2)$
- $(r+1, c-2)$
- $(r+1, c-1)$
- $(r+1, c)$
- $(r+1, c+1)$
- $(r+1, c+2)$
- $(r+2, c-2)$
- $(r+2, c-1)$
- $(r+2, c)$
- $(r+2, c+1)$
- $(r+2, c+2)$
Neighborhood Joint Probability

\[ N^2(r, c) = \{(u, v) | (u, v) = (r, c) + (i, j), i, j \in \{-2, -1, 0, 1, 2\}\} \]

\[ P(I(r, c), I(u, v) : (u, v) \in N^2(r, c)) = P(I(r, c)) \prod_{(u, v) \in N^1_2(r, c)} P_1(I(u, v) | I(r, c)) \prod_{(u, v) \in N^2_2(r, c)} P_2(I(u, v) | I(r, c)) \prod_{(u, v) \in N^3_2(r, c)} P_3(I(u, v) | I(r, c)) \prod_{(u, v) \in N^4_2(r, c)} P_4(I(u, v) | I(r, c)) \]

\[ J(r, c) = \log(P(I(u, v) : (u, v) \in N^2(r, c)))) = \log(P(I(r, c))) + \sum_{(u, v) \in N^1_2(r, c)} \log(P_1(I(u, v) | I(r, c))) + \sum_{(u, v) \in N^2_2(r, c)} \log(P_2(I(u, v) | I(r, c))) + \sum_{(u, v) \in N^3_2(r, c)} \log(P_3(I(u, v) | I(r, c))) + \sum_{(u, v) \in N^4_2(r, c)} \log(P_4(I(u, v) | I(r, c))) \]
A Texture Image
Visual Coherence is lost in gray level permuted image. Probability dependency is the same. White means the gray level configuration in a 5x5 window in the original image has a high joint probability.
White means the gray level configuration in a 5x5 window in the original image has a high joint probability.
When the scale of the texture and size of window are comparable, the joint probability is about the same for each window.
Within each uniform texture area, the joint probability in a window is nearly the same when the texture scale and the window size are similar.
If the texture scale is larger than the window size, the joint probability in each window will be low around boundaries.
A Gray Level Permuted Image: Brodatz

Visual Coherence is lost. Probability Dependency is the same.
Within each uniform texture area, the joint probability in a window is nearly the same when the texture scale and the window size are similar.
Within each uniform texture area, the joint probability in a window is nearly the same when the texture scale and the window size are similar.
Conditional Independences
Bound
The Dependencies
Semi-graphoid

**Definition**

Let \( I \) be an index set containing the indexes of all the random variables. Let \( G \) be a collection of triples each of whose components are subsets of the index set \( I \). We write \( A \perp B \mid C \) if and only if the triple \((A, B, C)\) is in \( G \).

\( G \) is called a **Semi-graphoid** if and only if

- **Mutual Exclusivity:** \((A, B, C)\) is in \( G \) implies \( A \cap B = \emptyset, A \cap C = \emptyset, B \cap C = \emptyset \)
- **Symmetry:** \( A \perp B \mid C \) if and only if \( B \perp A \mid C \)
- **Decomposition:** \( A \perp B \cup D \mid C \) implies \( A \perp B \mid C \)
- **Weak Union:** \( A \perp B \cup C \mid D \) implies \( A \perp B \mid C \cup D \)
- **Contraction:** \( A \perp B \mid C \cup D \) and \( A \perp C \mid D \), imply \( A \perp B \cup C \mid D \)


Judea Pearl and Azaria Paz, “Graphoids: A Graph-Based Logic for Reasoning About Relevance Relations”, University of California, Los Angeles, Computer Science Department, CSD-850038, 1985.
Graphoid

Definition

Let $I$ be an index set containing the indexes of all the random variables. Let $G$ be a collection of triples each of whose components are subsets of the index set $I$. We write $A \perp B \mid C$ if and only if the triple $(A, B, C) \in G$. $G$ is called a **Graphoid** if and only if

- **Mutual Exclusivity:** $(A, B, C) \in G$ implies
  - $A \cap B = \emptyset$, $A \cap C = \emptyset$, $B \cap C = \emptyset$

- **Symmetry:** $A \perp B \mid C$ if and only if $B \perp A \mid C$

- **Decomposition:** $A \perp B \cup D \mid C$ implies $A \perp B \mid C$

- **Weak Union:** $A \perp B \cup C \mid D$ implies $A \perp B \mid C \cup D$

- **Contraction:** $A \perp B \mid C \cup D$ and $A \perp C \mid D$, imply $A \perp B \cup C \mid D$

- **Intersection:** $A \perp B \mid C \cup D$ and $A \perp C \mid B \cup D$ imply $A \perp B \cup C \mid D$
$I$ is an index set of all the random variables

$P(X_i : i \in I) > 0$

$G$ is a collection of triples each of whose components are subsets of the index set $I$

$$G = \left\{ (A, B, C) \in \mathcal{P}(I)^3 \mid \begin{array}{l}
A, B \neq \emptyset \\
A, B, C \text{ are disjoint} \\
A \perp B \mid C
\end{array} \right\}$$

Then $G$ is a graphoid
A graph \((N, E)\) is called a **Conditional Independence Graph** of a random variable set \(X = \{X_1, \ldots, X_M\}\) if and only if \(N = \{1, \ldots, M\}\), the index set for the variables in \(X\), and

\[
E^c = \{\{i, j\} \mid X_i \perp\!\!\!\perp X_j \mid X - \{X_i, X_j\}\}
\]

\(\{i, j\}\) not in the edge set means \(X_i \perp\!\!\!\perp X_j \mid X - \{X_i, X_j\}\)

### Proposition

- $P(x) > 0$
- $G = (N, E)$ is a conditional independence graph
- $A, B, C$ are disjoint subsets of $N$
- $A, B, C \neq \emptyset$

**If $B$ separates $A$ from $C$, then $A \perp C \mid B$**

Proposition

- \((u, v), (a, b) \in N^2(r, c)\)
- \(L\) is the unique path from \((u, v)\) to \((a, b)\)
- \((m, n) \in L - \{(u, v), (a, b)\}\)
- \(I(u, v)\) is a random variable indexed by \((u, v)\)
- \(I(a, b)\) is a random variable indexed by \((a, b)\)
- \(I(m, n)\) is a random variable indexed by \((m, n)\)

Then

\[ I(u, v) \perp I(a, b) \mid I(m, n) \]
Theorem

If a graph $G$ is triangulated graph and $C_1, \ldots, C_K$ are the cliques of $G$ put in running intersection order with separators $S_2, \ldots, S_K$, then

$$S_k = C_k \bigcap \left( \bigcup_{i=1}^{k-1} C_i \right), \quad k = 2, \ldots, K$$

then

$$P(x_1, \ldots, x_N) = \frac{\prod_{k=1}^{K} P(x_i : i \in C_k)}{\prod_{k=2}^{K} P(x_i : i \in S_k)}$$

N-Tuple Method

- Developed For Printed Character Recognition
- Each character is contained in an image of $M \times N$ pixels
- Each pixel is a binary 1 or a binary 0
- Designed for table lookup hardware

N-Tuple Method
N-Tuple Method

- Small number of pixel positions are randomly selected
- Each of these positions has a binary 0 or a binary 1
- Concatenate all the binary values to form a binary number
- Use this number to form an address in memory
- Have as many memory arrays as character classes
- Have multiple sets of such randomly selected pixel positions
N-Tuple Method

- $M$ pattern sets of randomly selected pixel positions
- $K$ character classes
- $T_{mk}$ lookup table for pattern set $m$ and class $k$
- $T_{mk}(b_m)$ holds a binary 1 if a character in the training set of class $k$ has the binary number $b_m$ for the $m^{th}$ pattern set
- A printed character produces $M$ binary numbers $b_1, \ldots, b_M$
- Compute
  - $f_k = \min_{m=1}^{M} T_{mk}(b_m)$
  - $f_k = \sum_{m=1}^{M} T_{mk}(b_m)$
- Assign the character to unique class $c_k$, if there is one, for which $f_k > 0$ is highest
- Otherwise reserve decision
Each of the possible pixel positions is a variable
Let $X_1, \ldots, X_N$ be the $N$ variables
Let $L_n$ be the possible values variable $X_n$ can take
Let $R$ be the training set for one class

\[ R \subseteq \bigotimes_{n=1}^{N} L_n \]
If $I$ is an index set and $R \subseteq \times_{i \in I} L_i$, then we say $(I, R)$ is an Indexed N-ary Relation on the range sets indexed by $I$.
Relation Join

Definition

Let $I, J, K$ be index sets with $K = I \cup J$. Let $R \subset \bigtimes_{i \in I} L_i$ and $S \subset \bigtimes_{j \in J} L_j$. Then the Relation Join of $(I, R)$ with $(J, S)$ is denoted by $(I, R) \otimes (J, S) = (K, T)$ where

$$T = \{ t \in \bigtimes_{k \in K} L_k \mid \pi_I(K, t) \in (I, R) \text{ and } \pi_J(K, t) \in (J, S) \}$$
The set of measurement tuples that could be assigned to a class $c$, is the relation join of the tables associated with class $c$.

**Theorem**

$$\{([N], x) | \pi_{J_m}([N], x) \in (J_m, T_{mc}), m = 1, \ldots, M\} = \otimes_{m=1}^{M}(J_m, T_{mc})$$

- $(J_m, T_{mc})$ is an indexed relation
- Defined on index set $J_m$ associated with pattern set $m$
- $T_{mc}$ Contains all binary tuples from pattern set $m$ that were class $c$
N-Tuple Method

- $M$ pattern sets of randomly selected pixel positions
- $K$ character classes
- $T_{mk}$ lookup table for pattern set $m$ and class $k$
- $T_{mk}(b_m)$ holds the fraction of times a character in the training set of class $k$ has the binary number $b_m$ for the $m^{th}$ pattern set
- A printed character produces $M$ binary numbers $b_1, \ldots, b_M$
- Compute
  \[ f_k = \min_{m=1}^{M} T_{mk}(b_m) \]
  \[ f_k = \sum_{m=1}^{M} \log T_{mk}(b_m) \]
- For (1) Assign the character to unique class $c_k$, if there is one, for which $f_k > 0$
- For (2) Assign the character to unique class $c_k$, if there is one, for which $f_k > -\infty$
- Otherwise reserve decision
\[ f_k = \sum_{m=1}^{M} \log T_{mk}(b_m) \]

This is equivalent to

\[ \log P_c(x_1, \ldots, x_N) = \sum_{k=1}^{K} \log P_c(x_i : i \in C_k) \]

The class conditional independence assumption assumed here is surely wrong.
\[ P_c(x_1, \ldots, x_N) = \frac{\prod_{k=1}^{K} P_c(x_i : i \in C_k)}{\prod_{k=2}^{K} P_c(x_i : i \in S_k)} \]

\[ = P_c(x_i : i \in C_1) \prod_{k=2}^{K} P_c(x_i : i \in C_k - S_k | x_j : j \in S_k) \]

\[ \log P_c(x_1, \ldots, x_N) = \log P_c(x_i : i \in C_1) + \sum_{k=2}^{K} \log P_c(x_i : i \in C_k - S_k | x_j : j \in S_k) \]
Conditional No Influences
Bound
The Dependencies
Conditional No Influence

- $L_1 = \{a_1, a_2, a_3\}$
- $L_2 = \{c_1, c_2\}$
- $L_3 = \{b_1, b_2, b_3\}$
- $I = \{1, 2, 3\}$

$R \subset \bigotimes_{i \in I} L_i$

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$a_1$</td>
<td>$c_1$</td>
<td>$b_1$</td>
</tr>
<tr>
<td>2</td>
<td>$a_1$</td>
<td>$c_1$</td>
<td>$b_2$</td>
</tr>
<tr>
<td>3</td>
<td>$a_1$</td>
<td>$c_1$</td>
<td>$b_3$</td>
</tr>
<tr>
<td>4</td>
<td>$a_2$</td>
<td>$c_1$</td>
<td>$b_1$</td>
</tr>
<tr>
<td>5</td>
<td>$a_2$</td>
<td>$c_1$</td>
<td>$b_2$</td>
</tr>
<tr>
<td>6</td>
<td>$a_2$</td>
<td>$c_1$</td>
<td>$b_3$</td>
</tr>
<tr>
<td>7</td>
<td>$a_1$</td>
<td>$c_2$</td>
<td>$b_4$</td>
</tr>
<tr>
<td>8</td>
<td>$a_1$</td>
<td>$c_2$</td>
<td>$b_5$</td>
</tr>
<tr>
<td>9</td>
<td>$a_3$</td>
<td>$c_2$</td>
<td>$b_4$</td>
</tr>
<tr>
<td>10</td>
<td>$a_3$</td>
<td>$c_2$</td>
<td>$b_5$</td>
</tr>
</tbody>
</table>

1 has no influence on 3 given 2
\[ L_1 = \{a_1, a_2, a_3\} \]
\[ L_2 = \{c_1, c_2\} \]
\[ L_3 = \{b_1, b_2, b_3\} \]
\[ I = \{1, 2, 3\} \]
\[ J = \{1, 2\} \]
\[ K = \{2, 3\} \]

\[(I, R) = \pi_J(I, R) \otimes \pi_K(I, R)\]

\[ J - K \perp K - J \mid J \cap K \]

**Theorem**

Let \( G = \{(A, B, C) \in \mathcal{P}^3(I) \mid A \perp B \mid C\} \). Then \( G \) is a Semi-graphoid.

If a given relation is factored into a decomposition of $N$ factor relations, then grouping these factor relations into two possibly overlapping groups will also factor the given relation.

**Proposition**

Let $((M, R) = \bigotimes_{n=1}^{N} \pi_{M_n}(M, R)$ and $S \cup T = \{1, \ldots, N\}$. $S, T \neq \emptyset$. Then

$$\pi_{\cup_{s \in S} M_s}(M, R) \otimes \pi_{\cup_{t \in T} M_t}(M, R) = (M, R)$$

Corollary

Let \((M, R) = \bigotimes_{n=1}^{N} \pi_{M_n}(M, R)\). Define \(\mathcal{T} = \{ T \mid \text{for some } S \subset [N], S \neq \emptyset, T = \bigcup_{s \in S} M_s \}\), then \(U, V \in \mathcal{T}\) implies

\[ U - V \perp V - U \mid U \cap V \]

**Cliques and No Influence Pair of Sets**

\[ M = \{1, 2, 3, 4, 5, 6, 7, 8, 9\} \]

\[(M, R) = \otimes_{n=1}^{8} \pi_{M_n}(M, R)\]

\[ I = \{1, 2\}, \ J = \{8, 9\}, \ I \cap J = \emptyset \]

\[ M - (I \cup J) = \{3, 4, 5, 6, 7\} \]

\[ I \perp J \mid M - (I \cup J) \]

\[ I \cup (M - (I \cup J)) = M - J \]

\[ J \cup (M - (I \cup J)) = M - I \]

\[ (M, R) = \pi_{M-I}(M, R) \otimes \pi_{M-J}(M, R) \]

\[ I \times J \subseteq M \times M - \bigcup_{n=1}^{8} M_n \times M_n \]

\[ S = \{n \mid M_n \subseteq M - J\} = \{1, 2, 5, 6, 7\} \]

\[ T = \{n \mid M_n \subseteq M - I\} = \{3, 4, 6, 7, 8\} \]

\[ M - J = M_1 \cup M_2 \cup M_5 \cup M_6 \cup M_7 = \bigcup_{s \in S} M_s \]

\[ M - I = M_3 \cup M_4 \cup M_6 \cup M_7 \cup M_8 = \bigcup_{t \in T} M_t \]

\[ (o) \text{ Influence Graph} \]

\[ \begin{array}{|c|c|c|c|}
\hline
\text{Clique} & \text{Symbol} & \text{Clique} & \text{M – J} & \text{M – I} \\
\hline
M_1 & \{1, 2, 4\} & 1 & 0 \\
M_2 & \{2, 4, 5\} & 1 & 0 \\
M_3 & \{5, 6, 9\} & 0 & 1 \\
M_4 & \{5, 8, 9\} & 0 & 1 \\
M_5 & \{2, 3\} & 1 & 0 \\
M_6 & \{3, 6\} & 1 & 1 \\
M_7 & \{4, 7\} & 1 & 1 \\
M_8 & \{7, 8\} & 0 & 1 \\
\hline
\end{array} \]

\[ (p) \text{ Cliques} \]
Bach Choral BWV26

$R$ Set of 5-tuples obtained from Music Corpus

$I = \{1, 2, 3, 4, 5\}$

$J = \{3, 4, 5, 6, 7\}$

Construct the join $(K, S) = (I, R) \otimes (J, R)$

**Proposition**

If $(K, S) = (I, R) \otimes (J, R)$, $I - J \neq \emptyset$, and $J - I \neq \emptyset$

Then,

$I - J \perp J - I \mid I \cap J$

$\{1, 2\} \perp \{6, 7\} \mid \{3, 4, 5\}$
Joining The Tuples

\[ I_1 = \{1, 2, 3, 4, 5\} \]
\[ I_2 = \{3, 4, 5, 6, 7\} \]
\[ \vdots \]
\[ I_n = \{2n - 1, 2n, 2n + 1, 2n + 2, 2n + 3\} \]
\[ \vdots \]

- Construct \((K, S) = \otimes_{n=1}^{N} (I_n, R)\)
- Sample tuples from \((K, S)\) to listen to
Bach Generation By Relation Join

- 202 Bach Chorals Midi Files Taken from Music21
- Sequence of 5 Chords and Durations Constitute A Tuple
- Relation Join Requires Overlap of 3 Chords
- First Chord of Join Must be the I Chord of the Key
- Last Chord of the Join Must be a long duration I Chord of the Key

A Bach Choral

Random Sample Generated Bach Choral

- 4 Chords and Overlap 2 generate over 24 million sequences