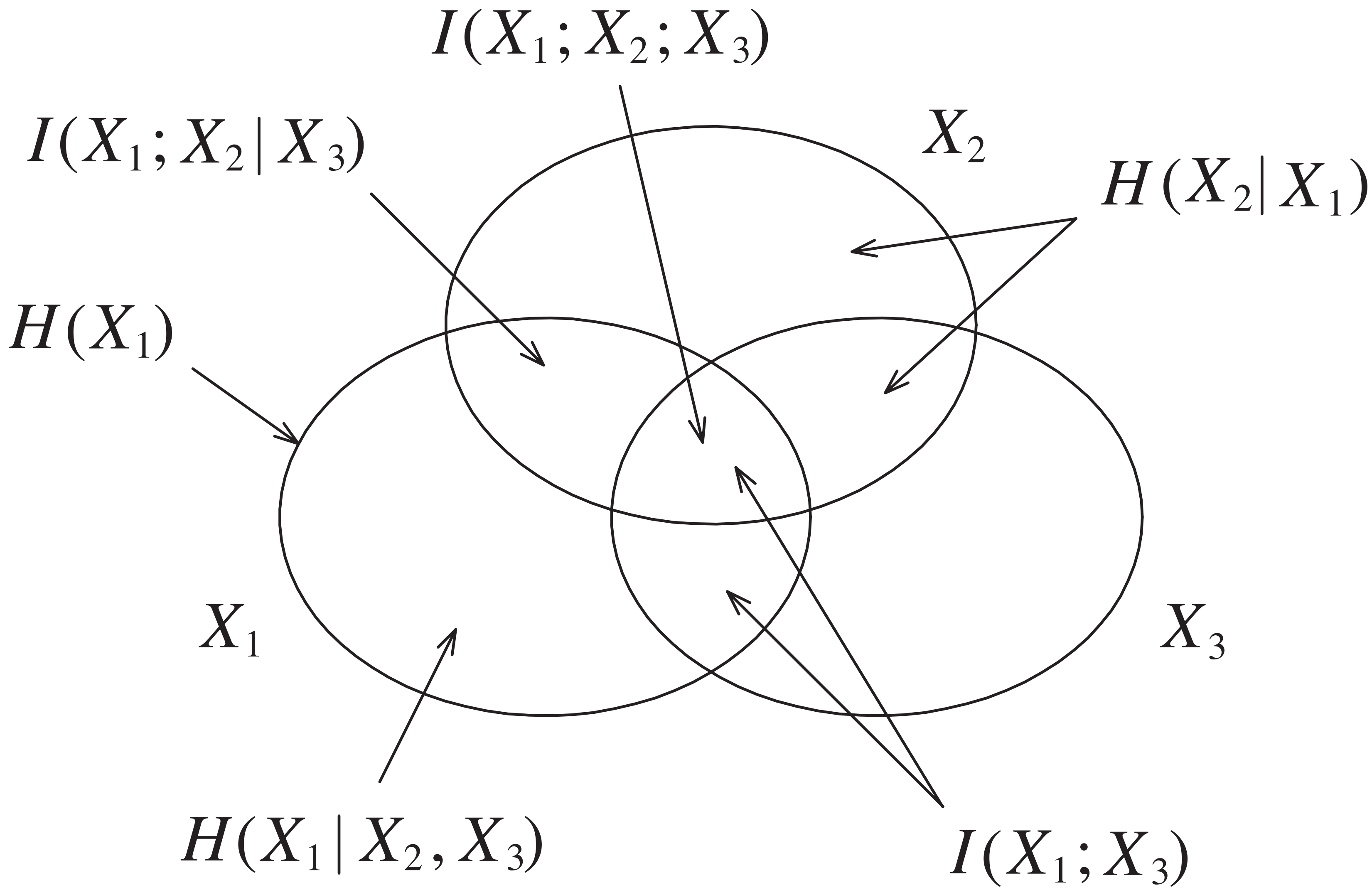
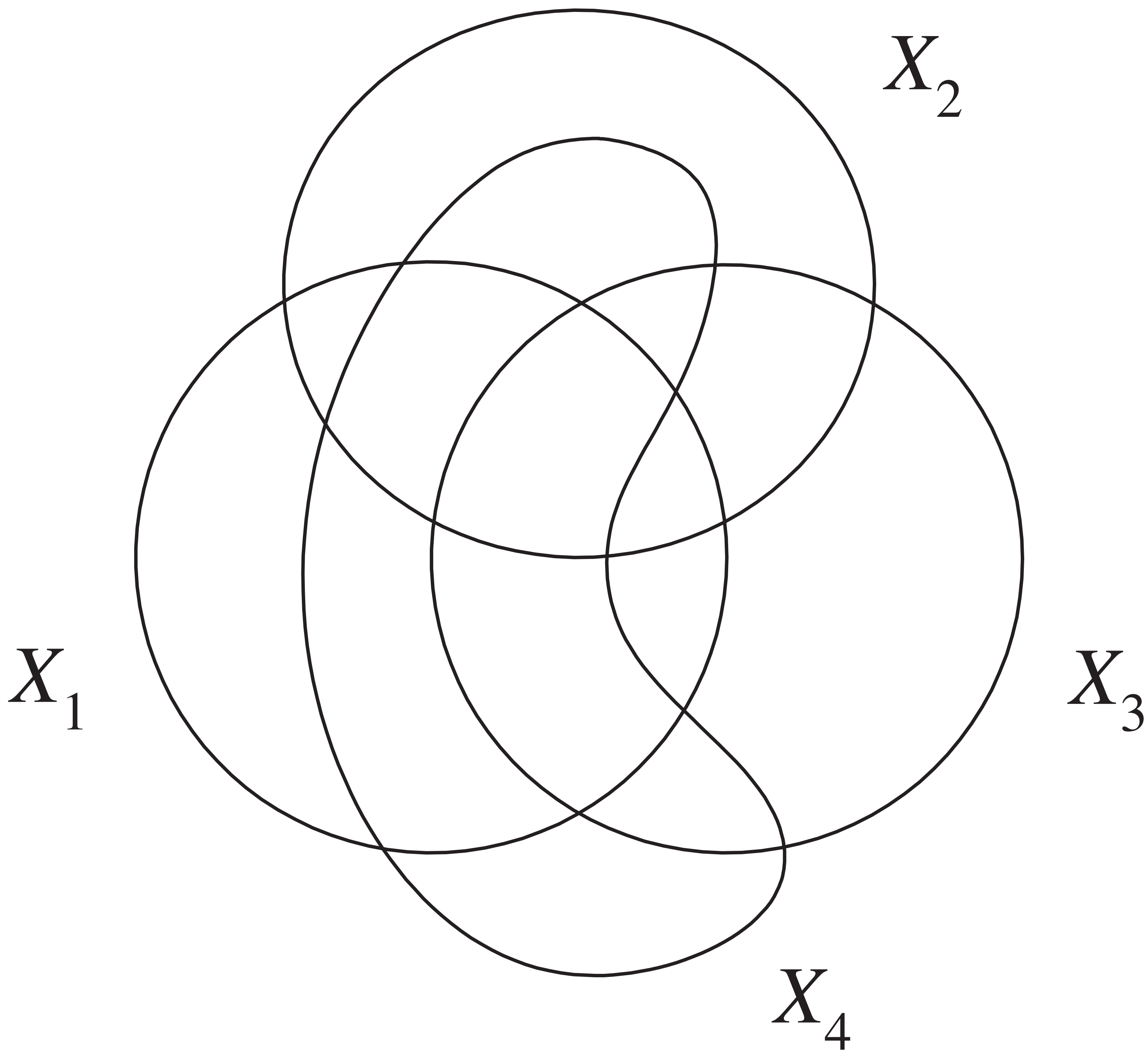
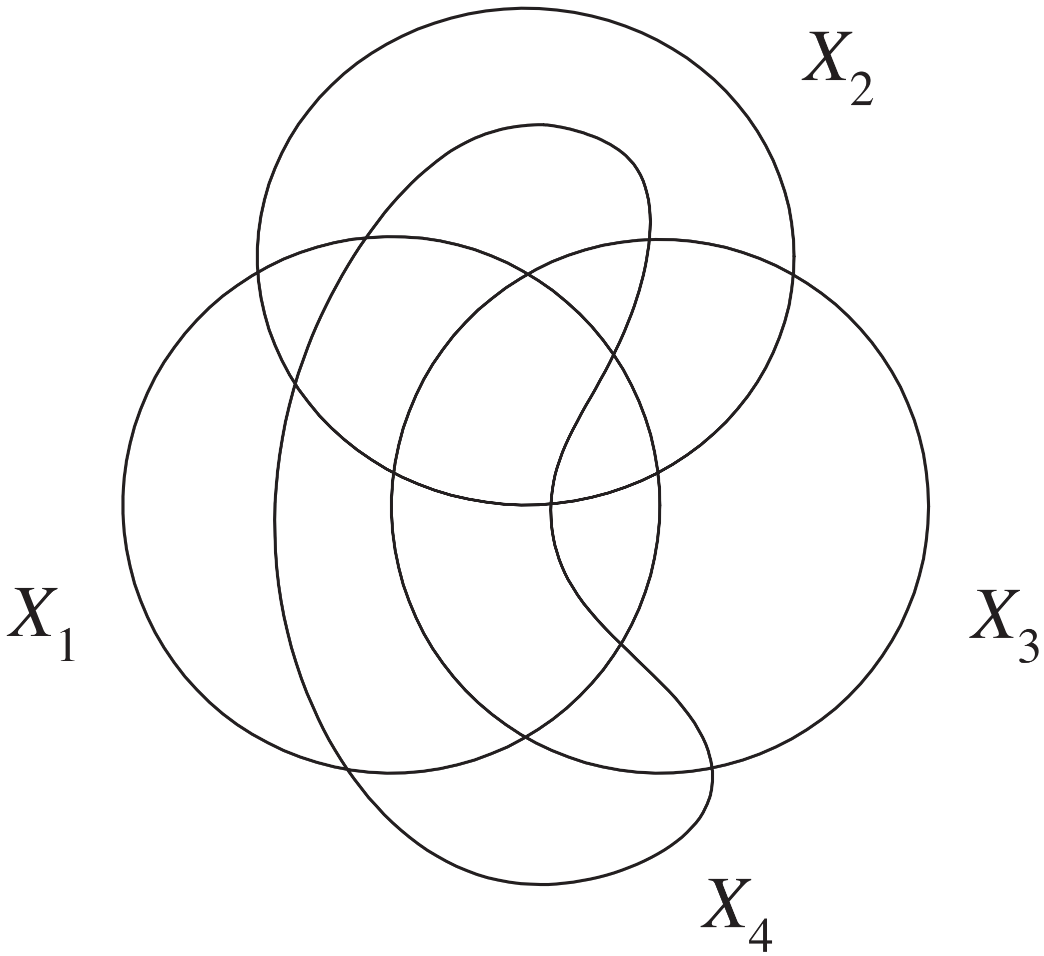
3.5 Information Diagrams

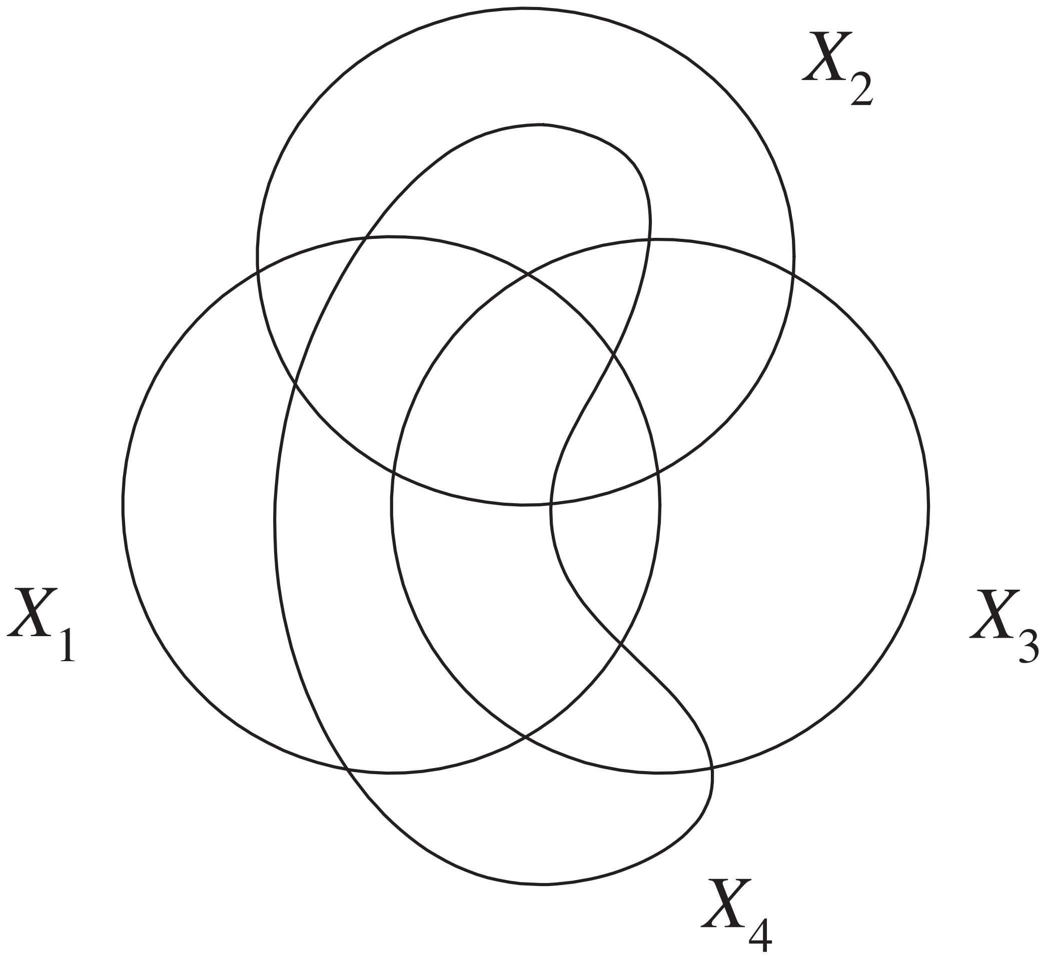

$I(X_1;X_2;X_3)$
$I(X_1;X_2|X_3)$
X_2
$H(X_2|X_1)$
$H(X_1)$
X_1
X_3
$H(X_1|X_2,X_3)$
$I(X_1;X_3)$



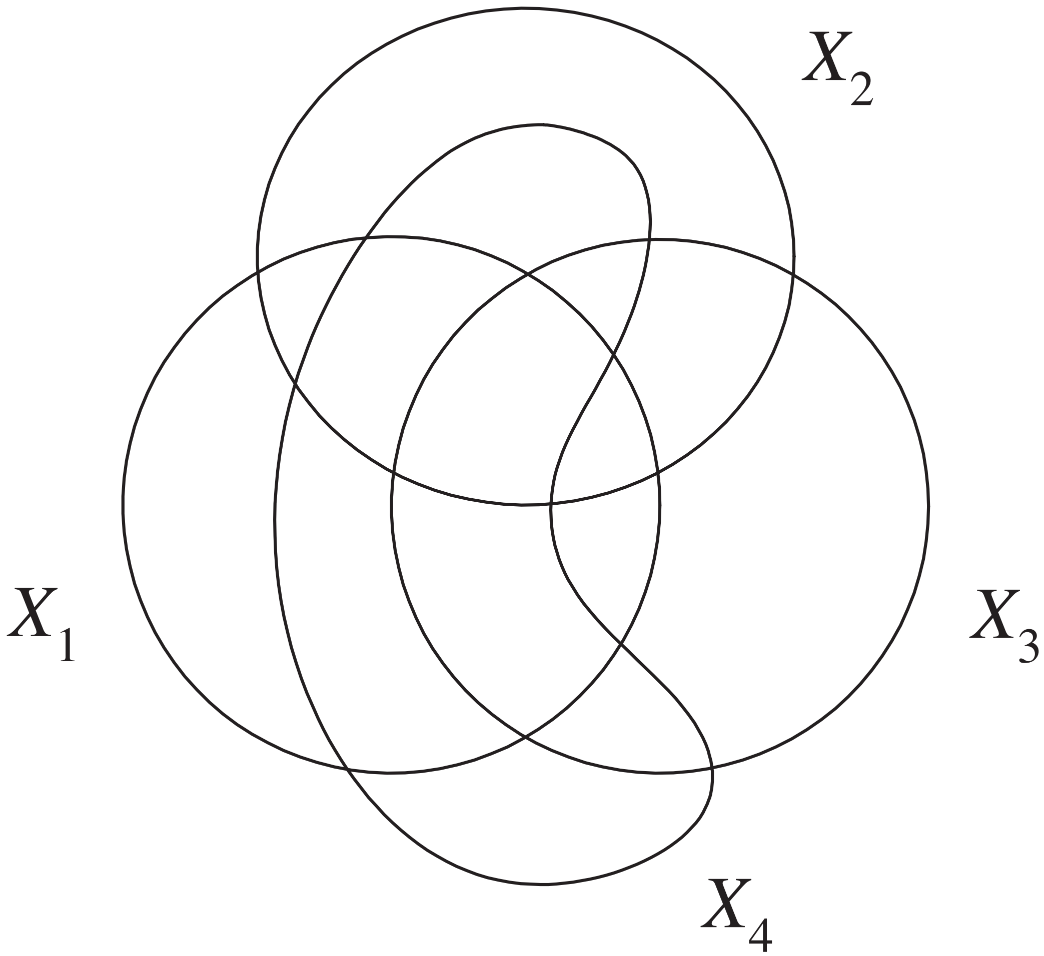
X_2
X_1
X_3
X_4



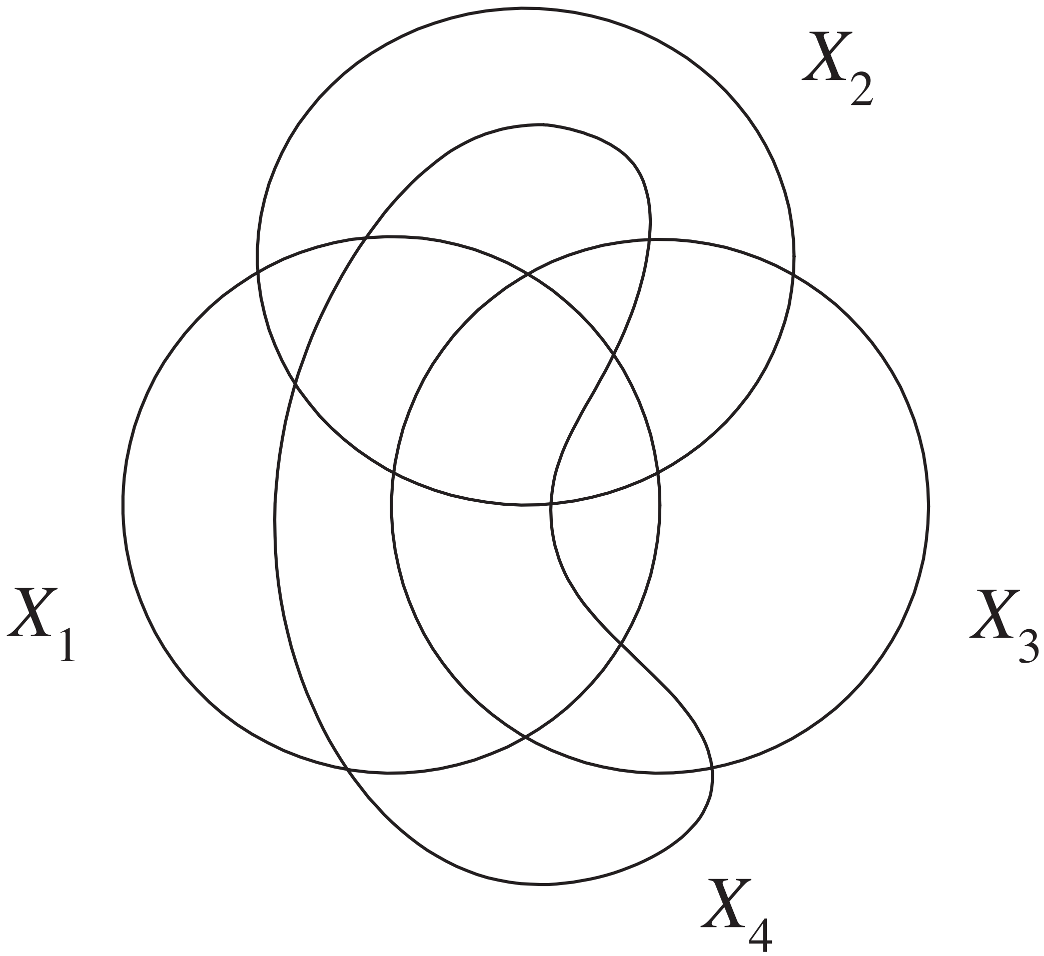
X_2
X_1
X_3
X_4



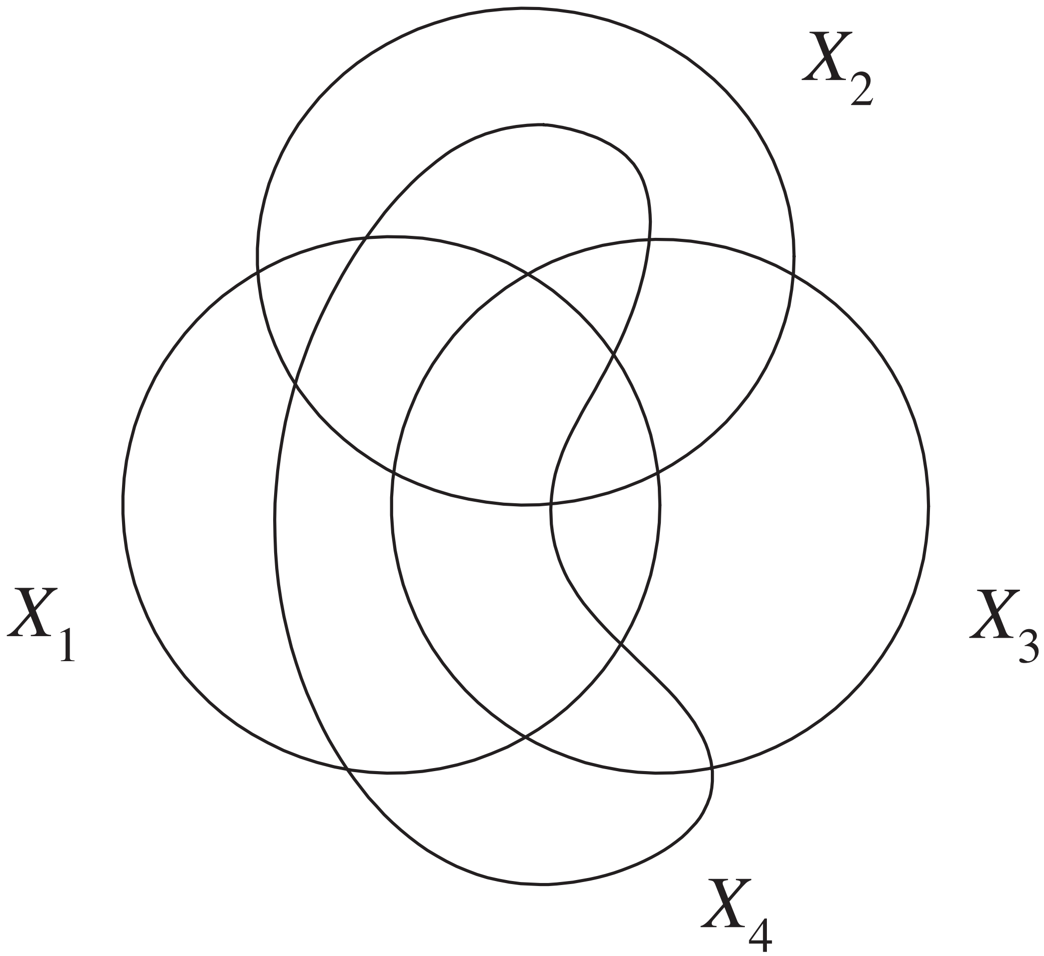
- This is a 2-dimensional representation of an information diagram for 4 r.v.'s.



- This is a 2-dimensional representation of an information diagram for 4 r.v.'s.
- First check that there are $2^4 - 1 = 15$ nonempty atoms.

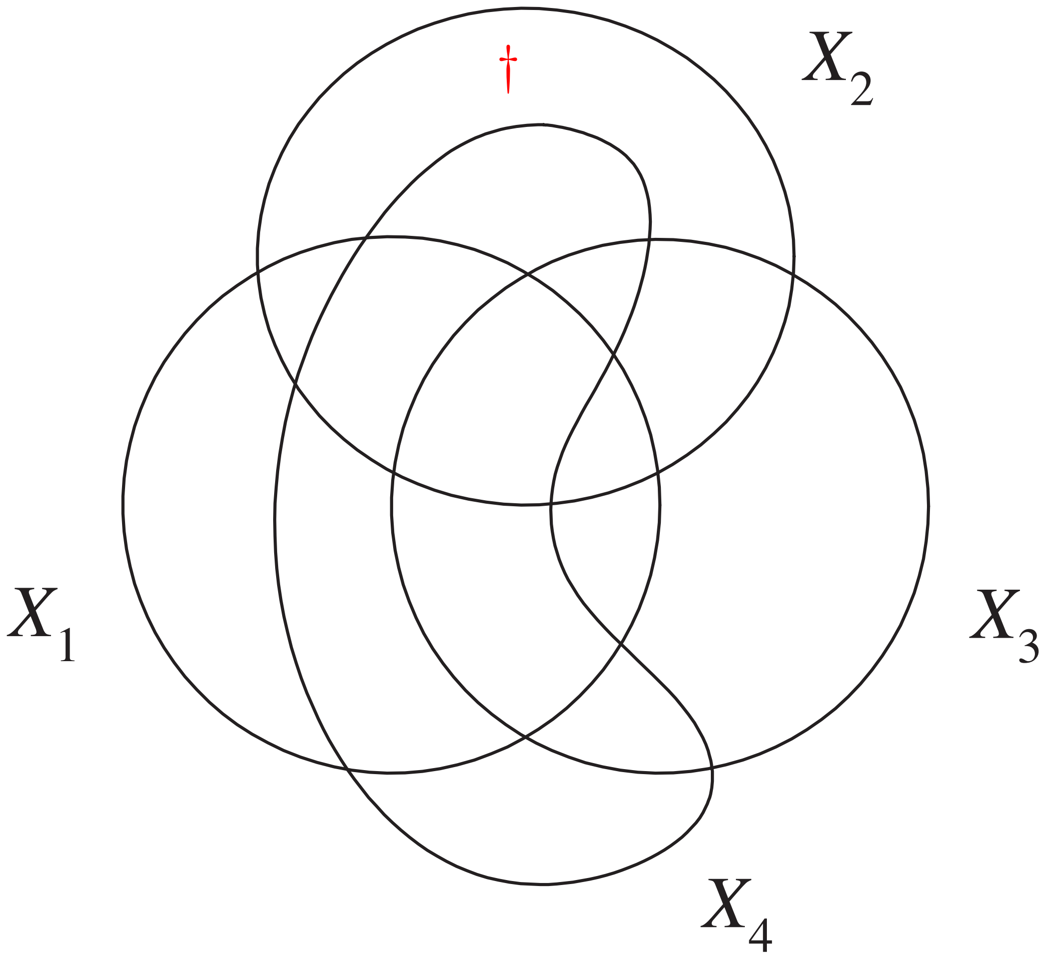


- This is a 2-dimensional representation of an information diagram for 4 r.v.'s.
- First check that there are $2^4 - 1 = 15$ nonempty atoms.
- This is a correct representation because the "kidney shape" representing X_4 splits each nonempty atom in $\mathcal{F}_3$ into 2 nonempty atoms in $\mathcal{F}_4$.



- This is a 2-dimensional representation of an information diagram for 4 r.v.'s.
- First check that there are $2^4 - 1 = 15$ nonempty atoms.
- This is a correct representation because the "kidney shape" representing X_4 splits each nonempty atom in $\mathcal{F}_3$ into 2 nonempty atoms in $\mathcal{F}_4$.
- For example, the atom

$$\underline{\tilde{X}_1^c \cap \tilde{X}_2 \cap \tilde{X}_3^c}$$

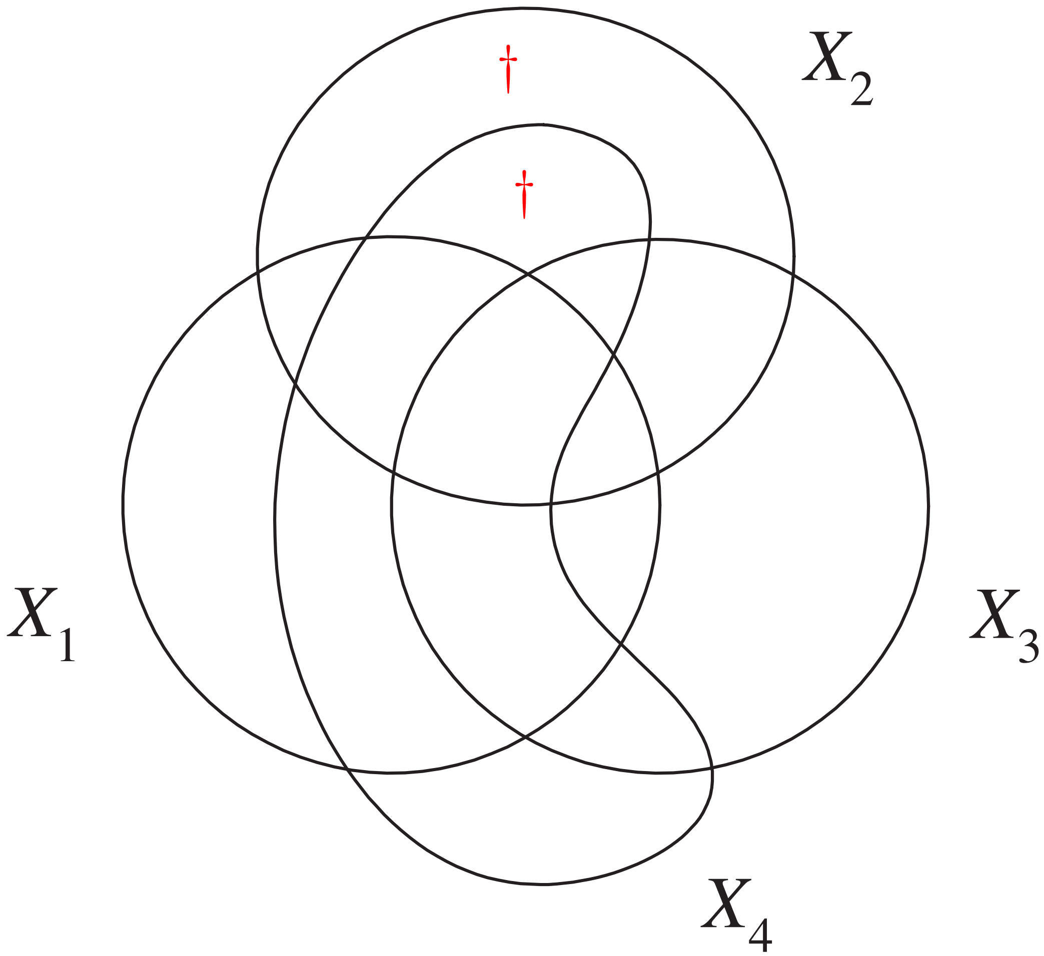


- This is a 2-dimensional representation of an information diagram for 4 r.v.'s.
- First check that there are $2^4 - 1 = 15$ nonempty atoms.
- This is a correct representation because the "kidney shape" representing X_4 splits each nonempty atom in $\mathcal{F}_3$ into 2 nonempty atoms in $\mathcal{F}_4$.
- For example, the atom

$$\underline{\tilde{X}_1^c \cap \tilde{X}_2 \cap \tilde{X}_3^c}$$

in $\mathcal{F}_3$ is split into the two atoms

$$\underline{\tilde{X}_1^c \cap \tilde{X}_2 \cap \tilde{X}_3^c} \cap \tilde{X}_4$$



- This is a 2-dimensional representation of an information diagram for 4 r.v.'s.
- First check that there are $2^4 - 1 = 15$ nonempty atoms.
- This is a correct representation because the "kidney shape" representing X_4 splits each nonempty atom in $\mathcal{F}_3$ into 2 nonempty atoms in $\mathcal{F}_4$.
- For example, the atom

$$\underline{\tilde{X}_1^c \cap \tilde{X}_2 \cap \tilde{X}_3^c}$$

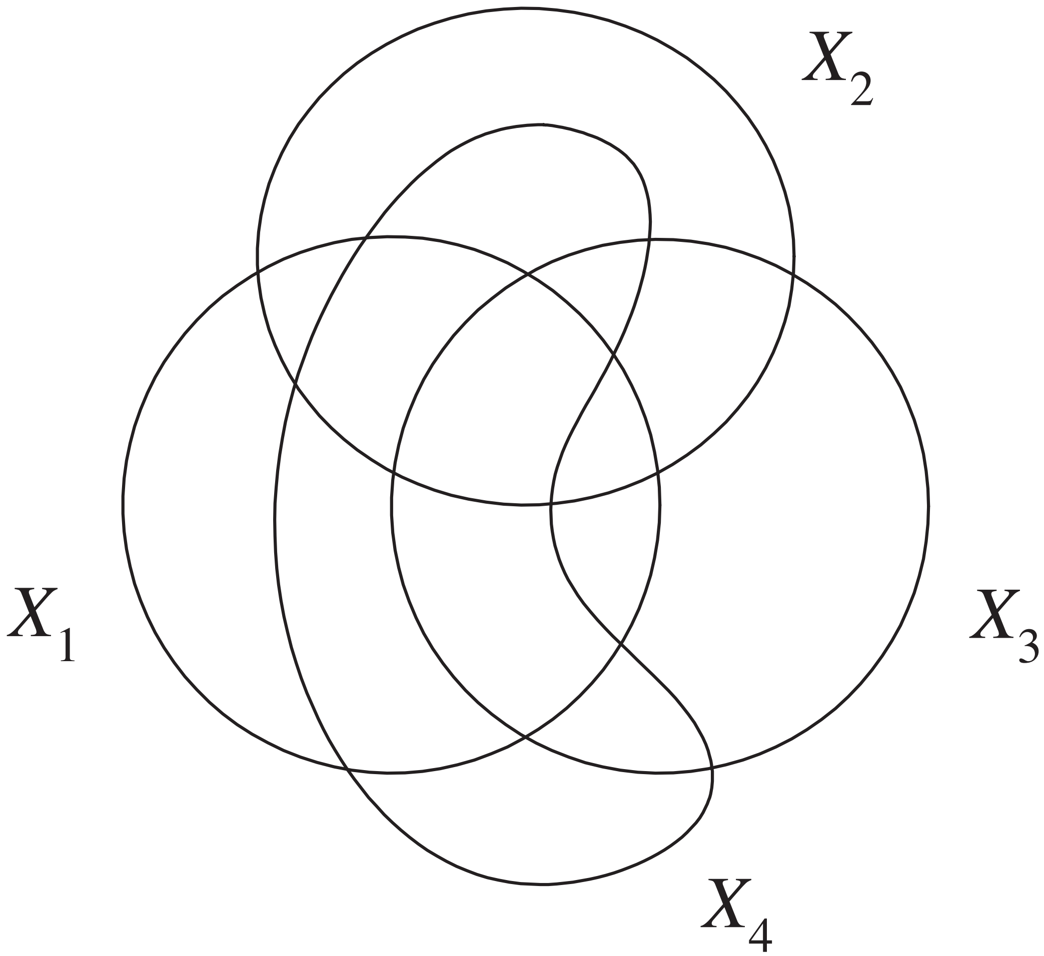
in $\mathcal{F}_3$ is split into the two atoms

$$\underline{\tilde{X}_1^c \cap \tilde{X}_2 \cap \tilde{X}_3^c} \cap \tilde{X}_4$$

and

$$\underline{\tilde{X}_1^c \cap \tilde{X}_2 \cap \tilde{X}_3^c} \cap \tilde{X}_4^c$$

in $\mathcal{F}_4$.



- This is a 2-dimensional representation of an information diagram for 4 r.v.'s.

- First check that there are $2^4 - 1 = 15$ nonempty atoms.

- This is a correct representation because the "kidney shape" representing X_4 splits each nonempty atom in $\mathcal{F}_3$ into 2 nonempty atoms in $\mathcal{F}_4$.

- For example, the atom

$$\tilde{X}_1^c \cap \tilde{X}_2 \cap \tilde{X}_3^c$$

in $\mathcal{F}_3$ is split into the two atoms

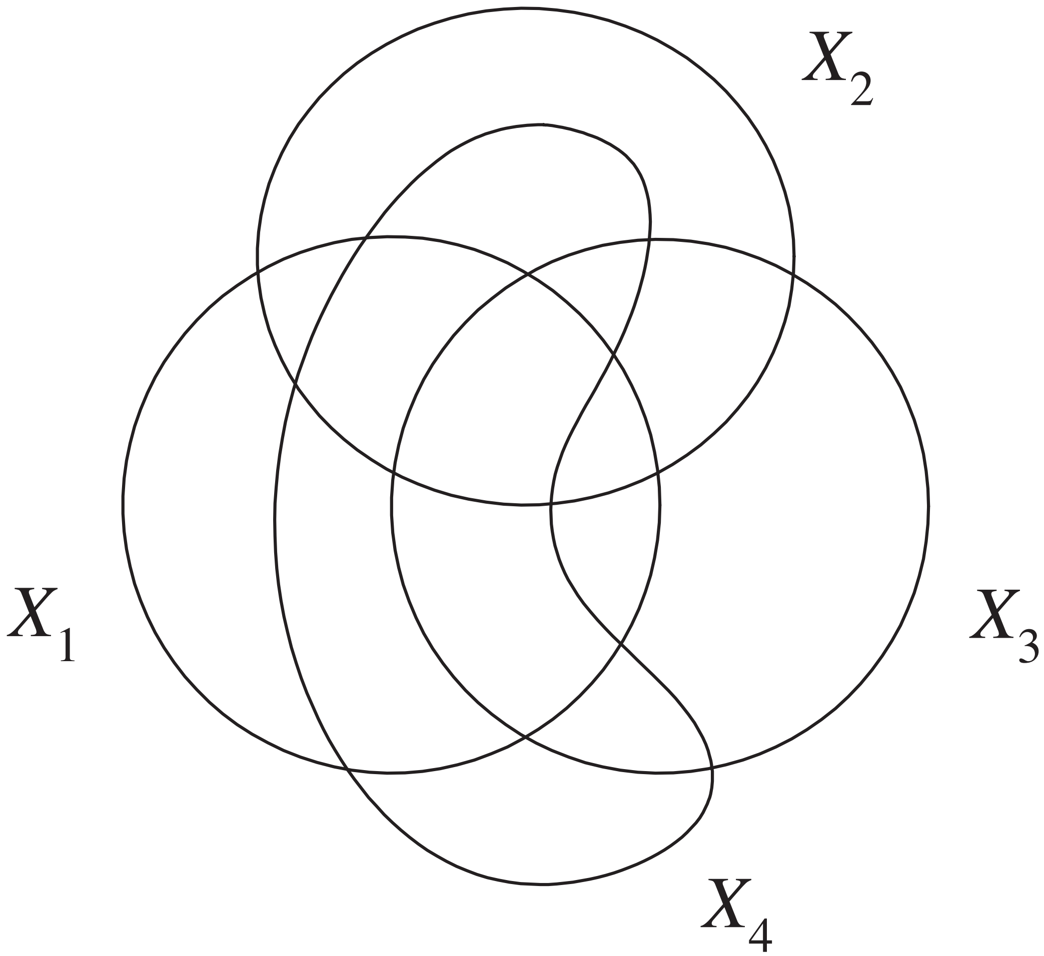
$$\tilde{X}_1^c \cap \tilde{X}_2 \cap \tilde{X}_3^c \cap \tilde{X}_4$$

and

$$\tilde{X}_1^c \cap \tilde{X}_2 \cap \tilde{X}_3^c \cap \tilde{X}_4^c$$

in $\mathcal{F}_4$.

- Exercise: Check that all the other 6 atoms of $\mathcal{F}_3$ are split into two atoms in $\mathcal{F}_4$.



- This is a 2-dimensional representation of an information diagram for 4 r.v.'s.

- First check that there are $2^4 - 1 = 15$ nonempty atoms.

- This is a correct representation because the "kidney shape" representing X_4 splits each nonempty atom in $\mathcal{F}_3$ into 2 nonempty atoms in $\mathcal{F}_4$.

- For example, the atom

$$\tilde{X}_1^c \cap \tilde{X}_2 \cap \tilde{X}_3^c$$

in $\mathcal{F}_3$ is split into the two atoms

$$\tilde{X}_1^c \cap \tilde{X}_2 \cap \tilde{X}_3^c \cap \tilde{X}_4$$

and

$$\tilde{X}_1^c \cap \tilde{X}_2 \cap \tilde{X}_3^c \cap \tilde{X}_4^c$$

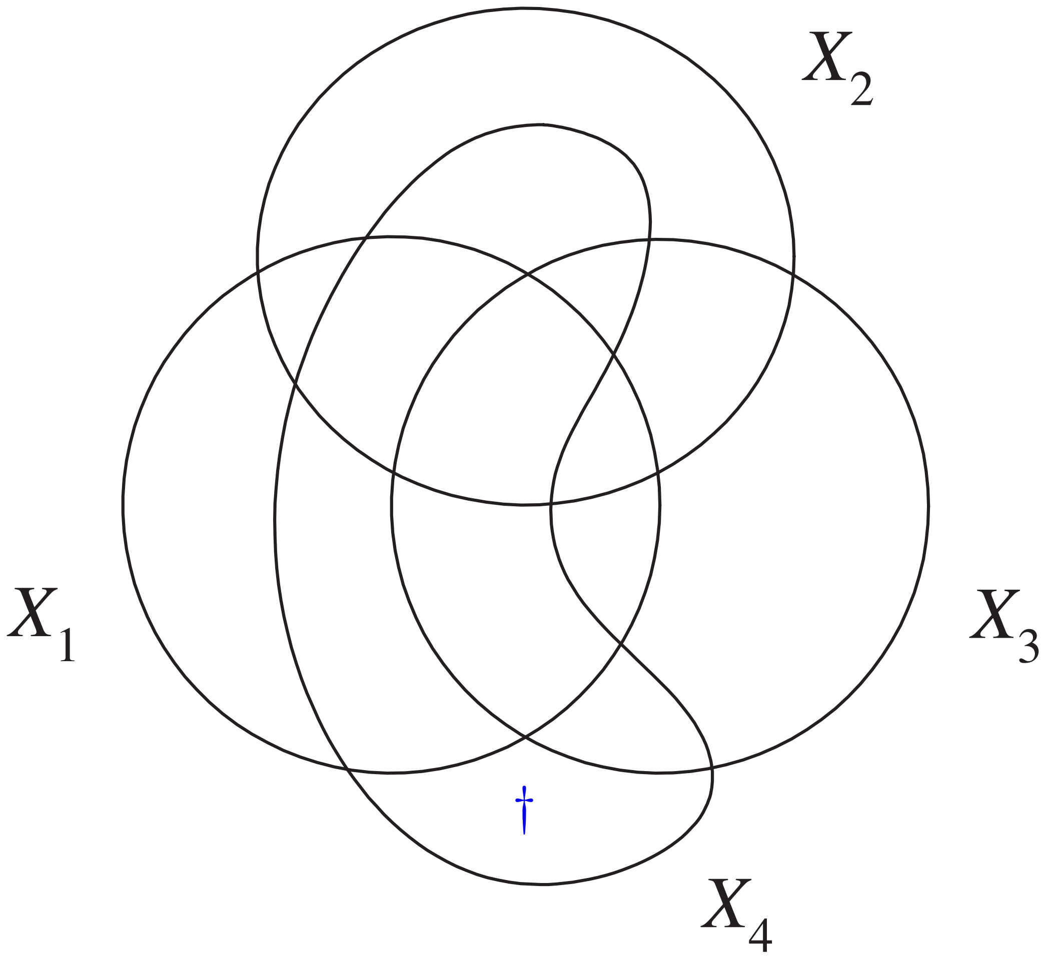
in $\mathcal{F}_4$.

- Exercise: Check that all the other 6 atoms of $\mathcal{F}_3$ are split into two atoms in $\mathcal{F}_4$.

- Also there is an extra nonempty atom

$$\tilde{X}_1^c \cap \tilde{X}_2^c \cap \tilde{X}_3^c \cap \tilde{X}_4$$

which is not in $\mathcal{F}_3$.



- This is a 2-dimensional representation of an information diagram for 4 r.v.'s.
- First check that there are $2^4 - 1 = 15$ nonempty atoms.
- This is a correct representation because the "kidney shape" representing X_4 splits each nonempty atom in $\mathcal{F}_3$ into 2 nonempty atoms in $\mathcal{F}_4$.
- For example, the atom

$$\tilde{X}_1^c \cap \tilde{X}_2 \cap \tilde{X}_3^c$$

in $\mathcal{F}_3$ is split into the two atoms

$$\tilde{X}_1^c \cap \tilde{X}_2 \cap \tilde{X}_3^c \cap \tilde{X}_4$$

and

$$\tilde{X}_1^c \cap \tilde{X}_2 \cap \tilde{X}_3^c \cap \tilde{X}_4^c$$

in $\mathcal{F}_4$.

- Exercise: Check that all the other 6 atoms of $\mathcal{F}_3$ are split into two atoms in $\mathcal{F}_4$.
- Also there is an extra nonempty atom

$$\tilde{X}_1^c \cap \tilde{X}_2^c \cap \tilde{X}_3^c \cap \tilde{X}_4$$

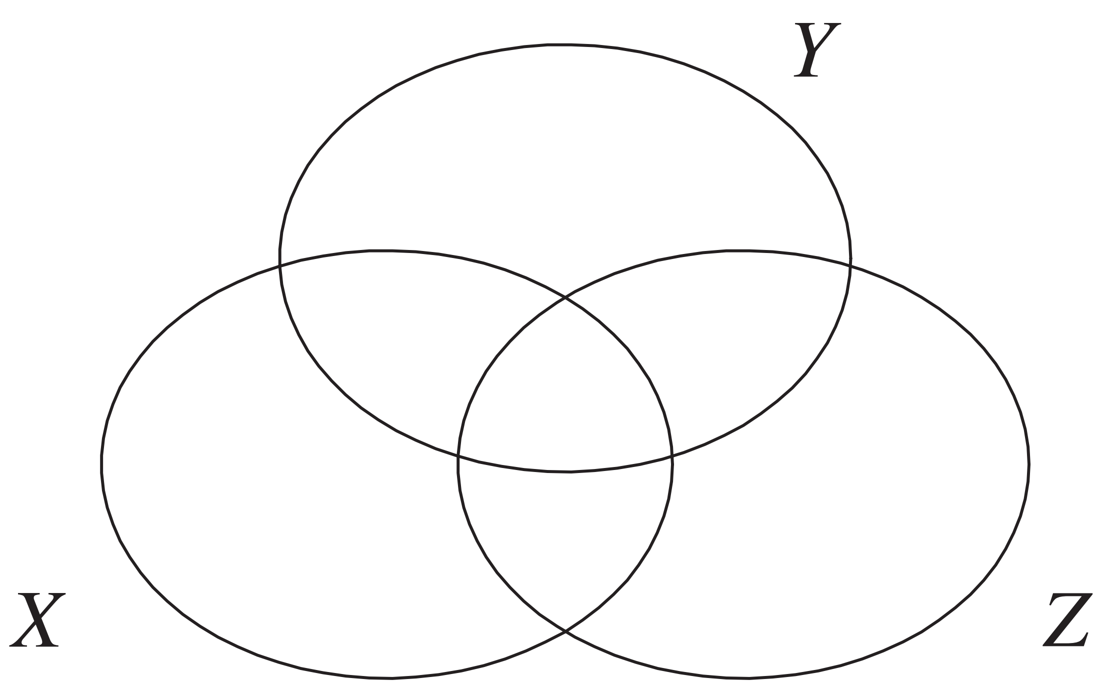
which is not in $\mathcal{F}_3$.

Theorem 3.11 If there is no constraint on $X_1, X_2, \cdots, X_n$, then μ^* can take any set of nonnegative values on the nonempty atoms of $\mathcal{F}_n$.

Idea

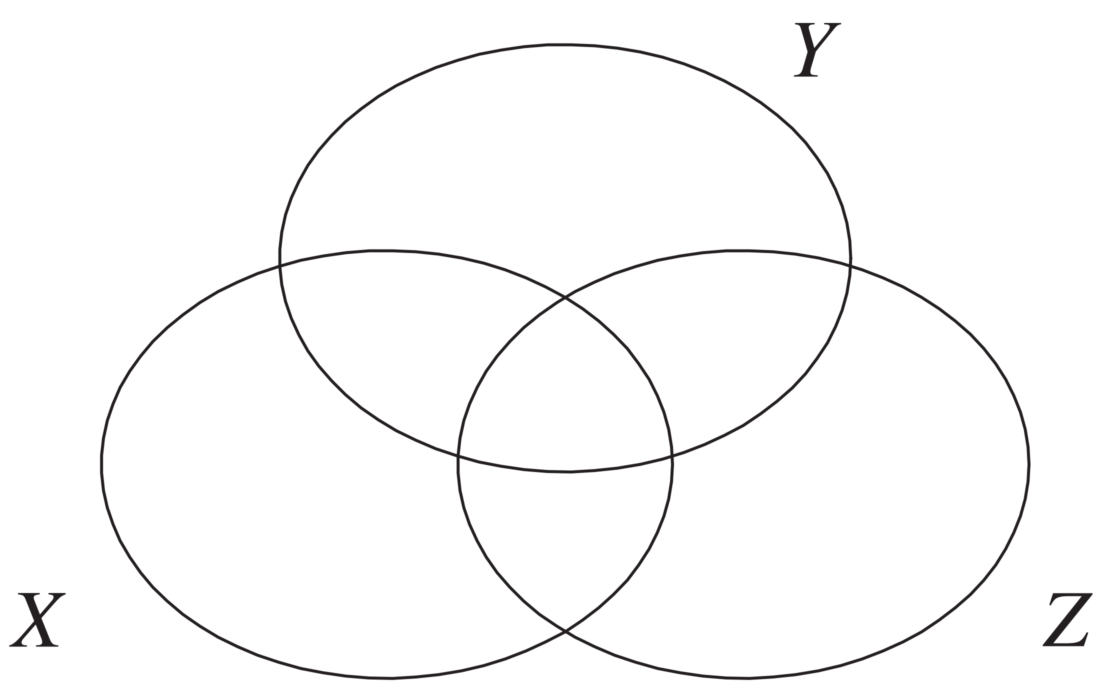
Theorem 3.11 If there is no constraint on $X_1, X_2, \cdots, X_n$, then μ^* can take any set of nonnegative values on the nonempty atoms of $\mathcal{F}_n$.

Idea



Theorem 3.11 If there is no constraint on $X_1, X_2, \cdots, X_n$, then μ^* can take any set of nonnegative values on the nonempty atoms of $\mathcal{F}_n$.

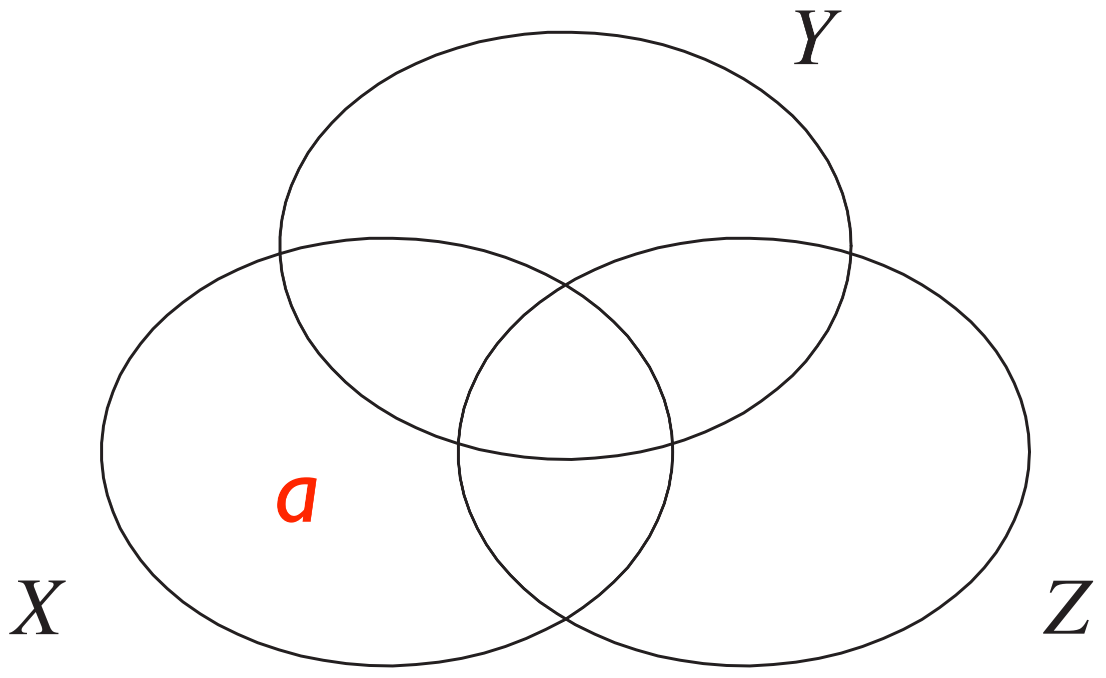
Idea



- Fill the nonempty atoms of the information diagram with arbitrary nonnegative numbers $a, b, \cdots, g$.

Theorem 3.11 If there is no constraint on $X_1, X_2, \cdots, X_n$, then μ^* can take any set of nonnegative values on the nonempty atoms of $\mathcal{F}_n$.

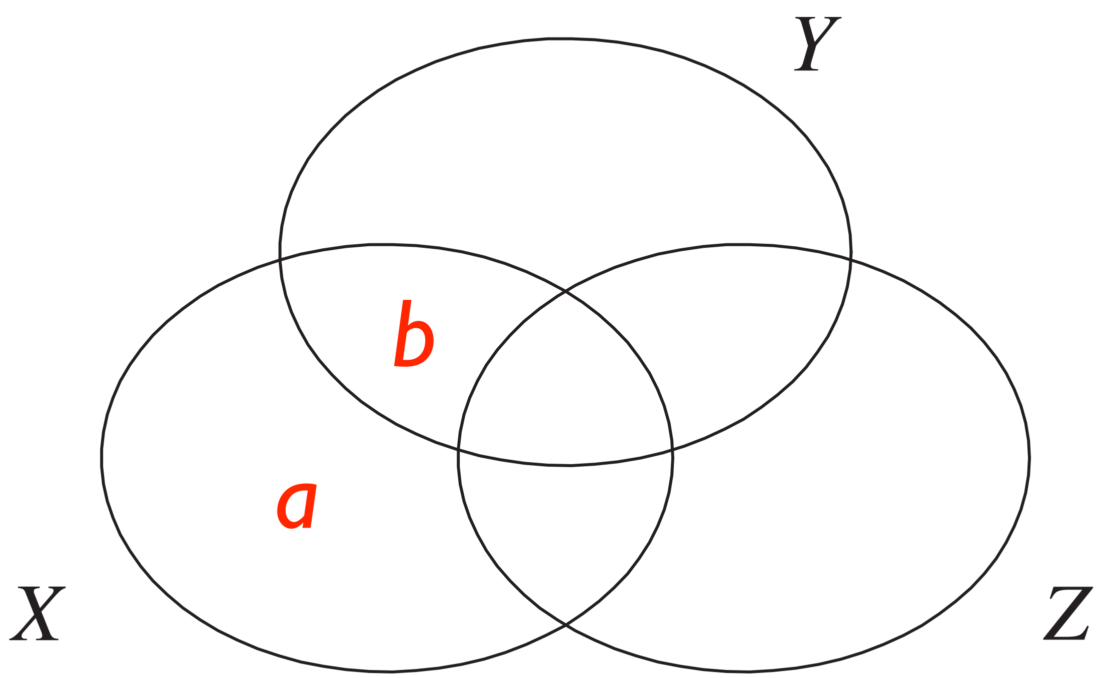
Idea



- Fill the nonempty atoms of the information diagram with arbitrary nonnegative numbers $a, b, \cdots, g$.

Theorem 3.11 If there is no constraint on $X_1, X_2, \cdots, X_n$, then μ^* can take any set of nonnegative values on the nonempty atoms of $\mathcal{F}_n$.

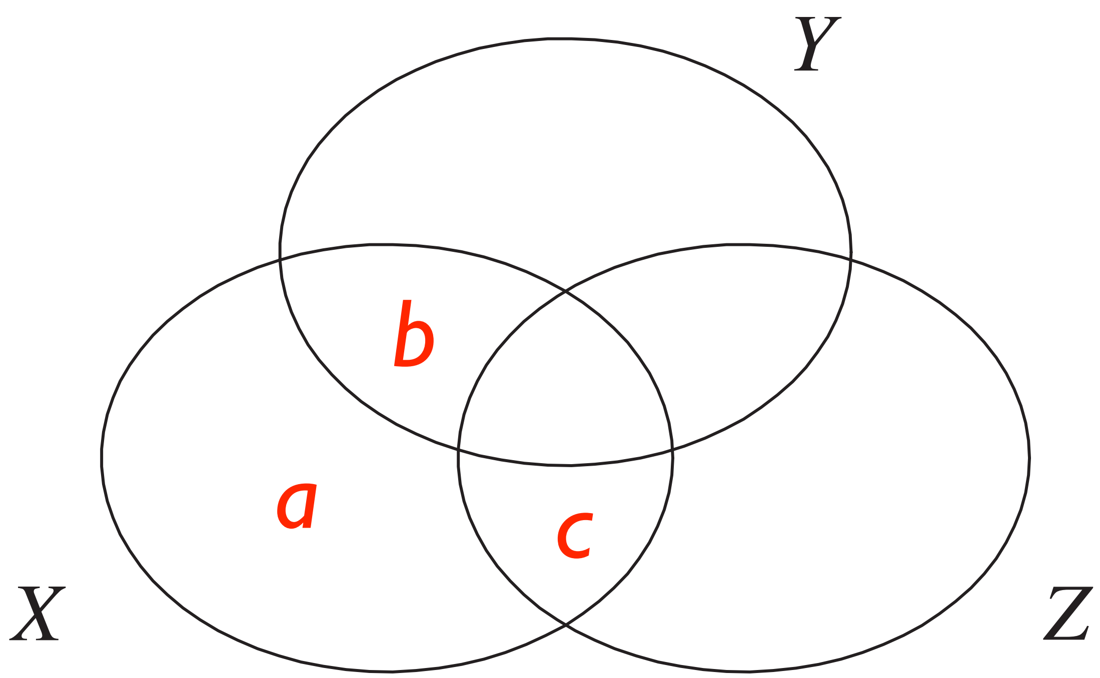
Idea



- Fill the nonempty atoms of the information diagram with arbitrary nonnegative numbers $a, b, \cdots, g$.

Theorem 3.11 If there is no constraint on $X_1, X_2, \cdots, X_n$, then μ^* can take any set of nonnegative values on the nonempty atoms of $\mathcal{F}_n$.

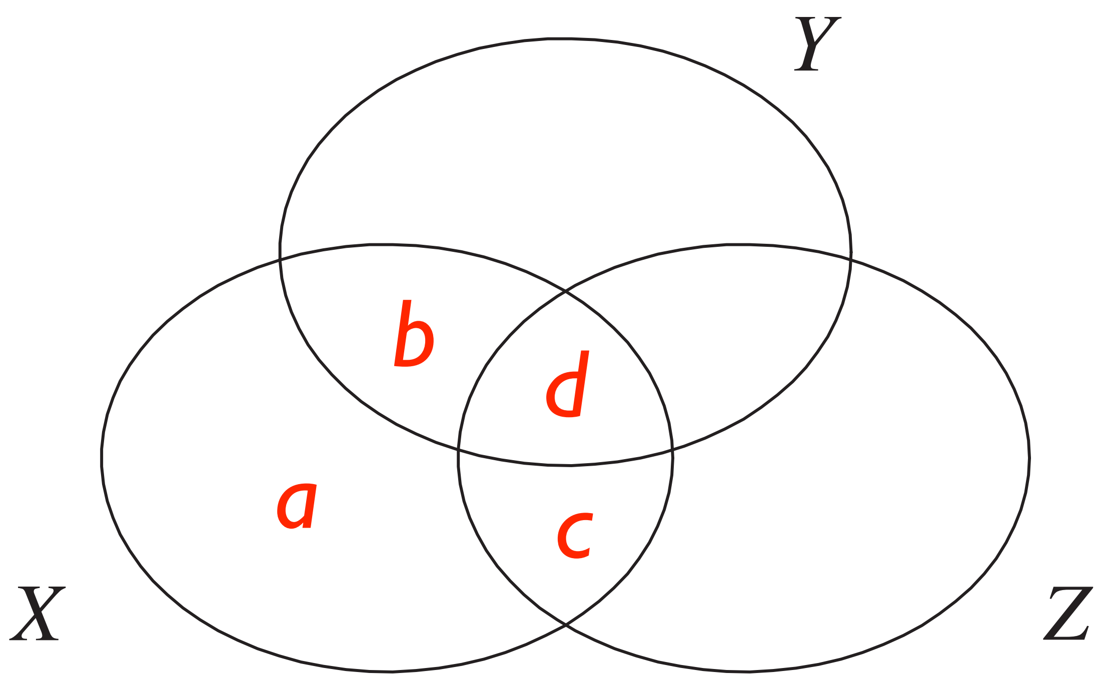
Idea



- Fill the nonempty atoms of the information diagram with arbitrary nonnegative numbers $a, b, \cdots, g$.

Theorem 3.11 If there is no constraint on $X_1, X_2, \cdots, X_n$, then μ^* can take any set of nonnegative values on the nonempty atoms of $\mathcal{F}_n$.

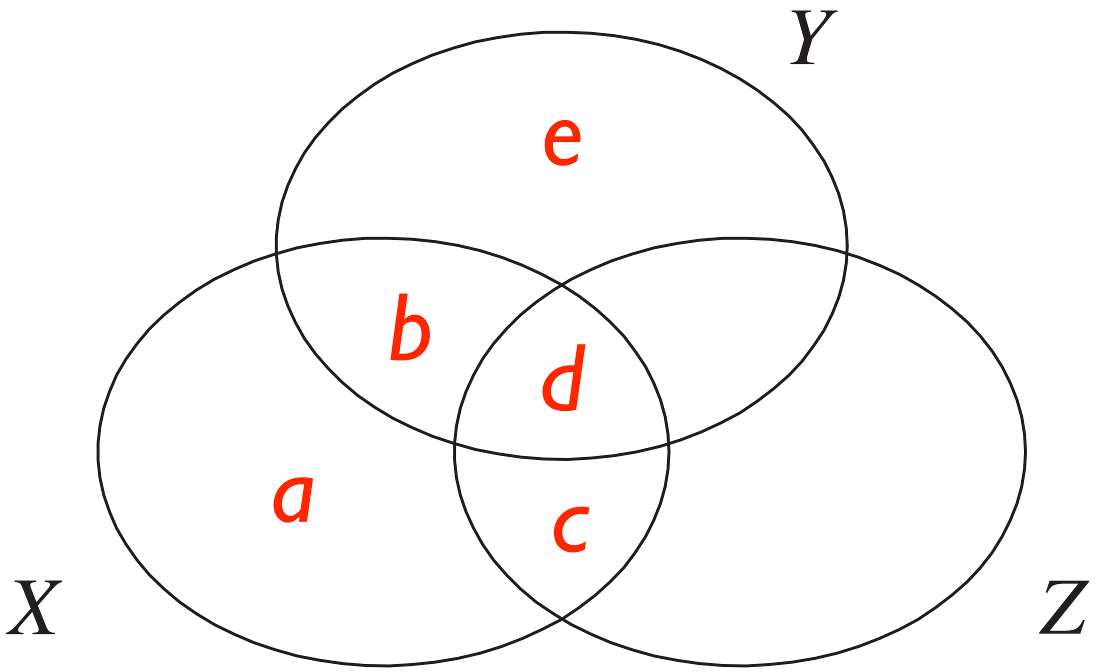
Idea



- Fill the nonempty atoms of the information diagram with arbitrary nonnegative numbers $a, b, \cdots, g$.

Theorem 3.11 If there is no constraint on $X_1, X_2, \cdots, X_n$, then μ^* can take any set of nonnegative values on the nonempty atoms of $\mathcal{F}_n$.

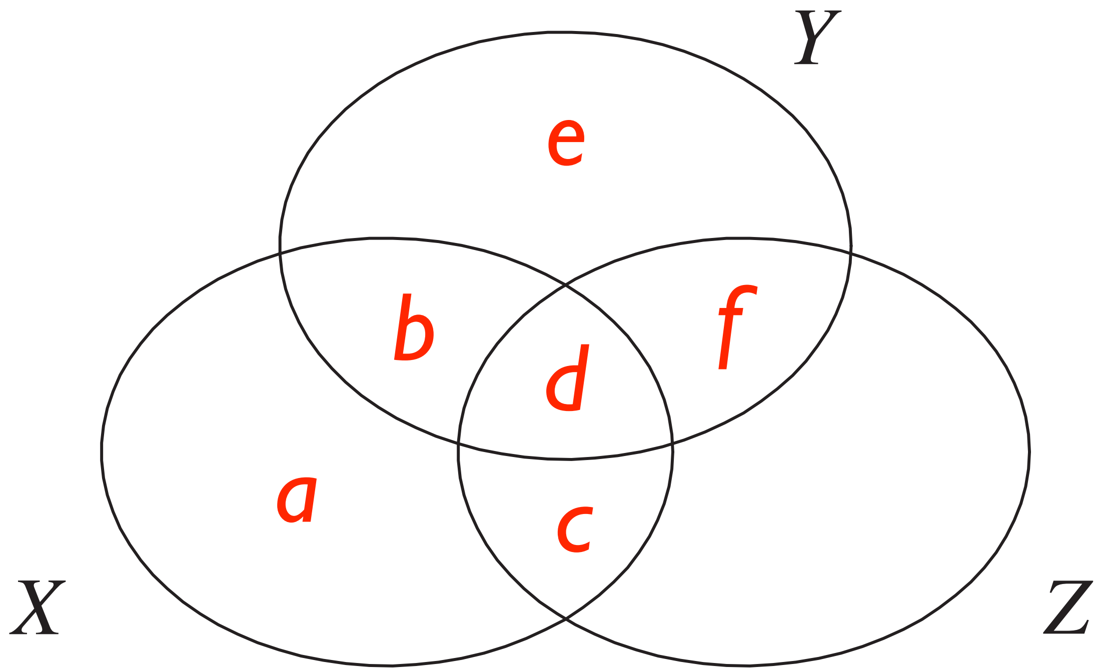
Idea



- Fill the nonempty atoms of the information diagram with arbitrary nonnegative numbers $a, b, \cdots, g$.

Theorem 3.11 If there is no constraint on $X_1, X_2, \cdots, X_n$, then μ^* can take any set of nonnegative values on the nonempty atoms of $\mathcal{F}_n$.

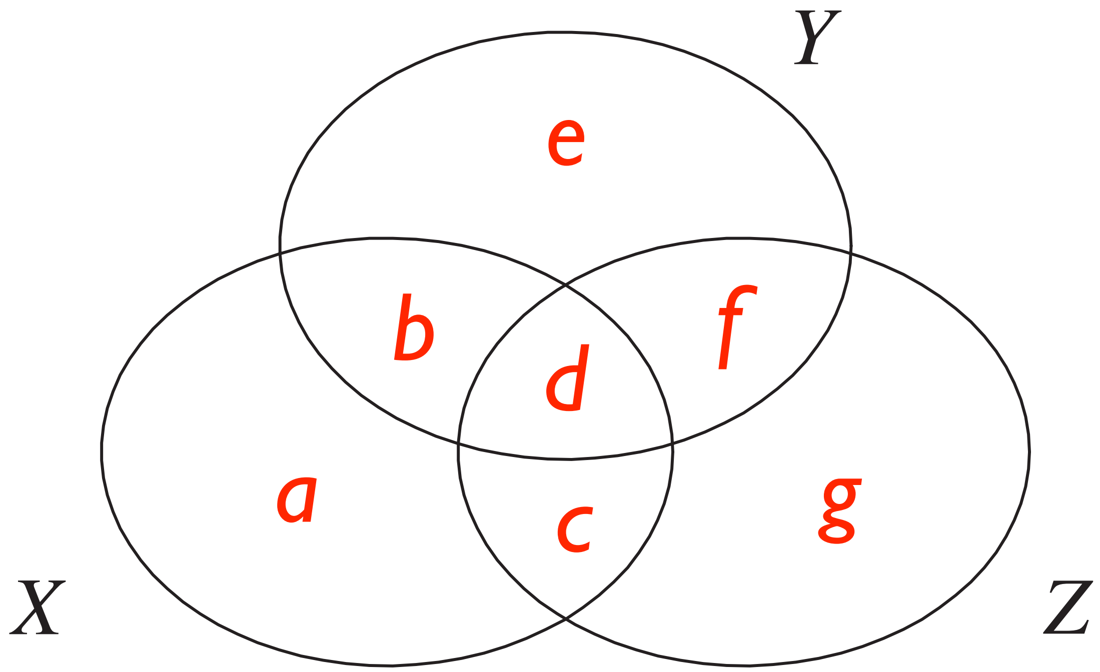
Idea



- Fill the nonempty atoms of the information diagram with arbitrary nonnegative numbers $a, b, \cdots, g$.

Theorem 3.11 If there is no constraint on $X_1, X_2, \cdots, X_n$, then μ^* can take any set of nonnegative values on the nonempty atoms of $\mathcal{F}_n$.

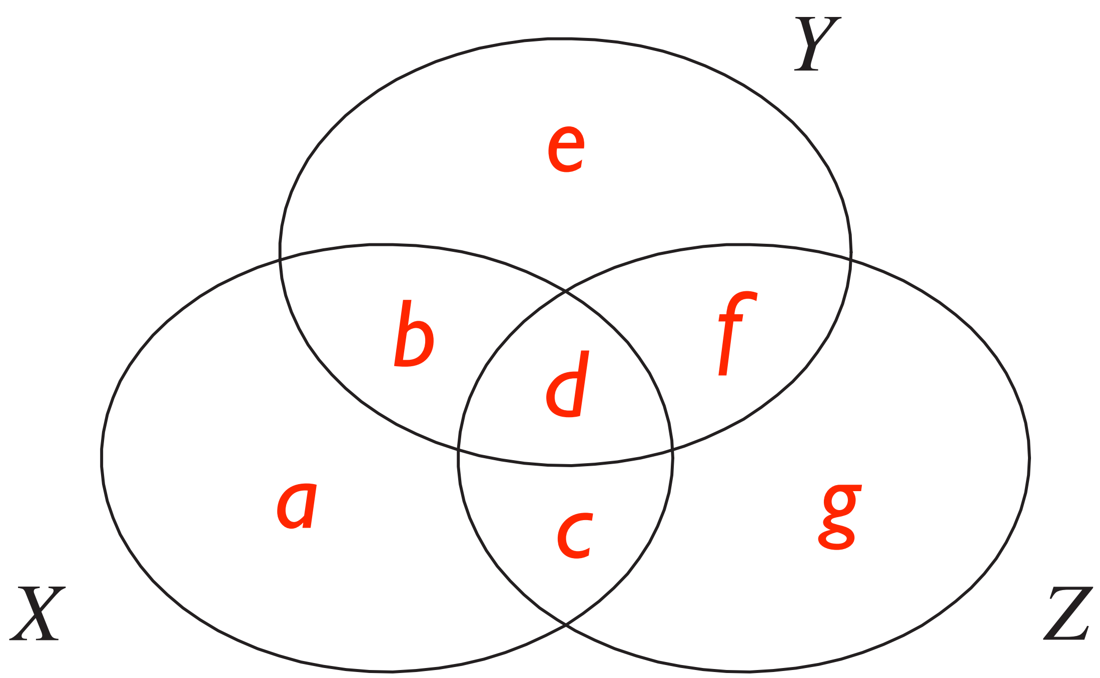
Idea



- Fill the nonempty atoms of the information diagram with arbitrary nonnegative numbers $a, b, \cdots, g$.

Theorem 3.11 If there is no constraint on $X_1, X_2, \cdots, X_n$, then μ^* can take any set of nonnegative values on the nonempty atoms of $\mathcal{F}_n$.

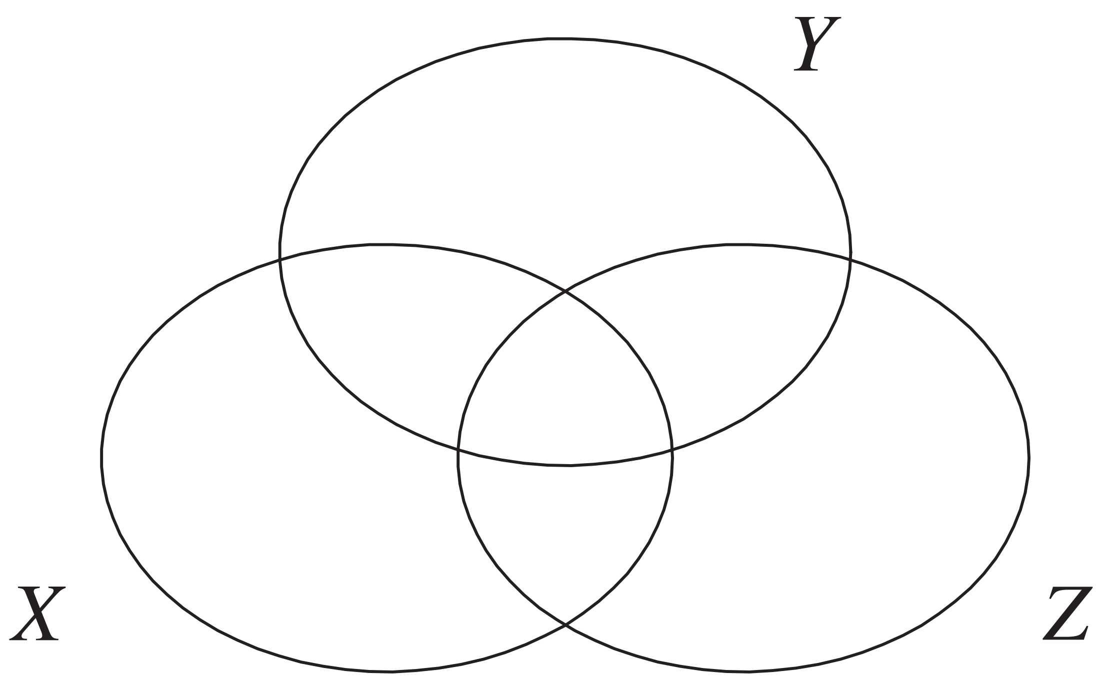
Idea



- Fill the nonempty atoms of the information diagram with arbitrary nonnegative numbers $a, b, \cdots, g$.
- Theorem 3.11 says that there exist r.v.'s X, Y, and Z whose I-measure μ^* is as shown.

Theorem 3.11 If there is no constraint on $X_1, X_2, \cdots, X_n$, then μ^* can take any set of nonnegative values on the nonempty atoms of $\mathcal{F}_n$.

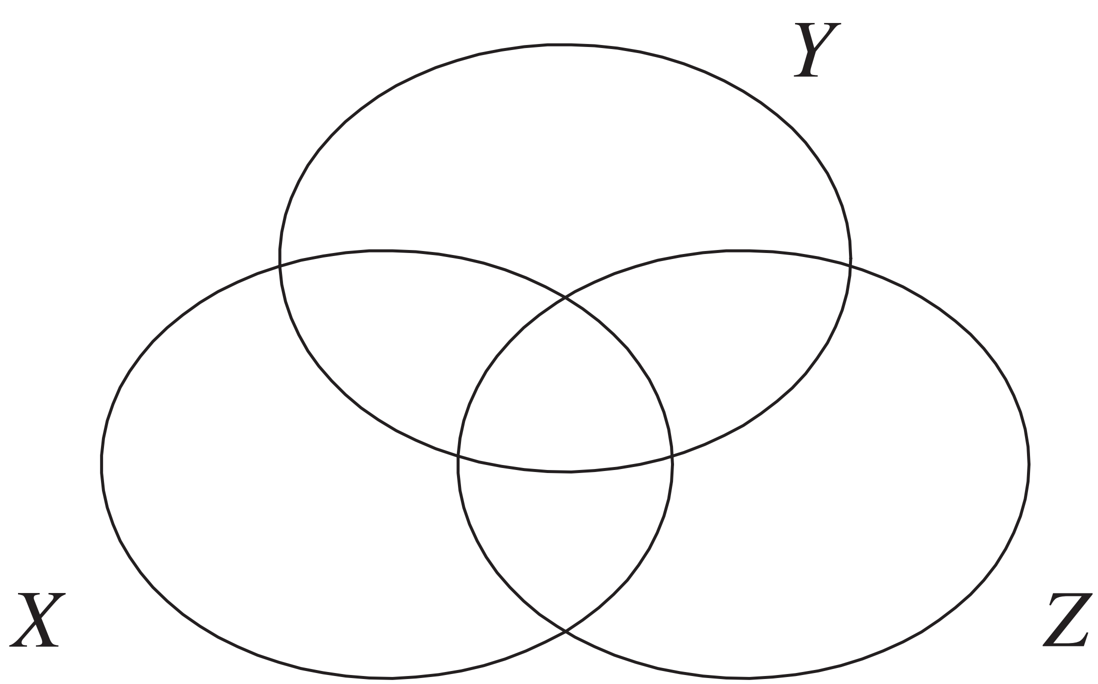
Idea



- Fill the nonempty atoms of the information diagram with arbitrary nonnegative numbers $a, b, \cdots, g$.
- Theorem 3.11 says that there exist r.v.'s X, Y, and Z whose I-measure μ^* is as shown.

Theorem 3.11 If there is no constraint on $X_1, X_2, \cdots, X_n$, then μ^* can take any set of nonnegative values on the nonempty atoms of $\mathcal{F}_n$.

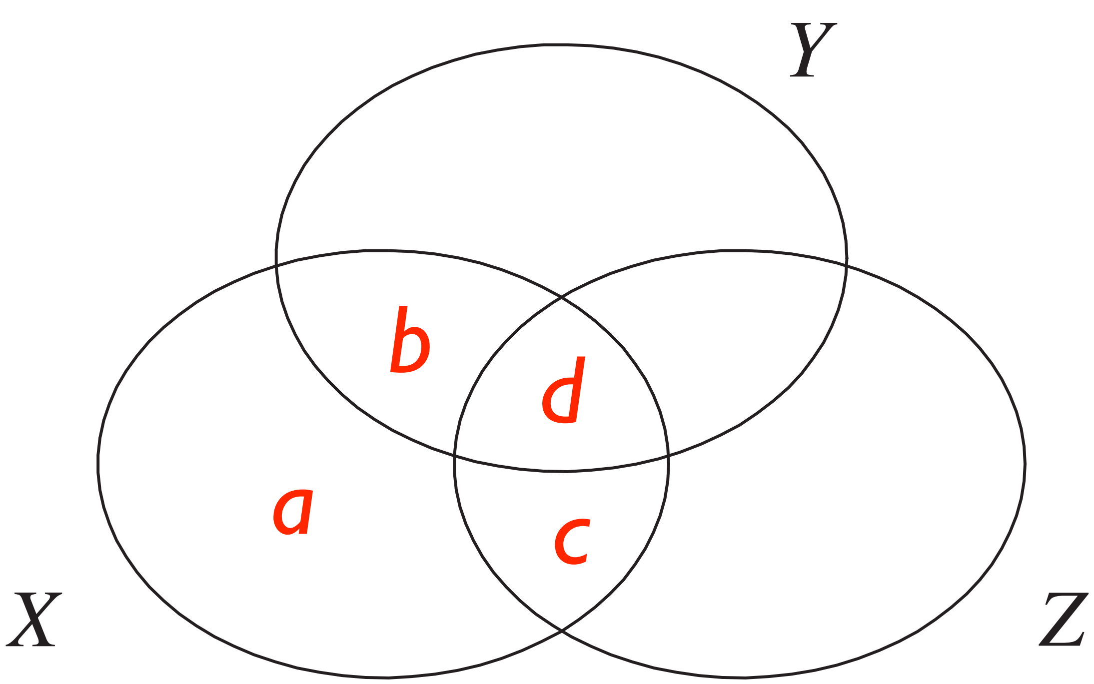
Idea



- Fill the nonempty atoms of the information diagram with arbitrary nonnegative numbers $a, b, \cdots, g$.
- Theorem 3.11 says that there exist r.v.'s X, Y, and Z whose I-measure μ^* is as shown.
- This can be seen by considering mutually independent r.v.'s $A, B, \cdots, G$ with entropies $a, b, \cdots, g$, respectively, and let

Theorem 3.11 If there is no constraint on $X_1, X_2, \cdots, X_n$, then μ^* can take any set of nonnegative values on the nonempty atoms of $\mathcal{F}_n$.

Idea

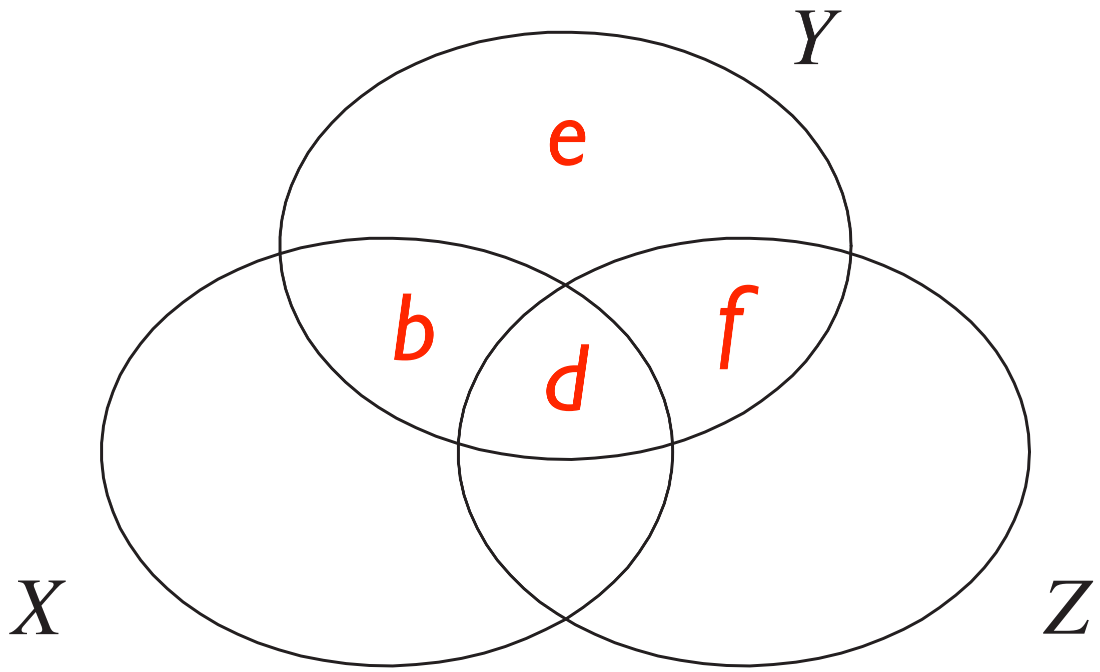


- Fill the nonempty atoms of the information diagram with arbitrary nonnegative numbers $a, b, \cdots, g$.
- Theorem 3.11 says that there exist r.v.'s X, Y, and Z whose I-measure μ^* is as shown.
- This can be seen by considering mutually independent r.v.'s $A, B, \cdots, G$ with entropies $a, b, \cdots, g$, respectively, and let

$$X = (A, B, C, D)$$

Theorem 3.11 If there is no constraint on $X_1, X_2, \cdots, X_n$, then μ^* can take any set of nonnegative values on the nonempty atoms of $\mathcal{F}_n$.

Idea

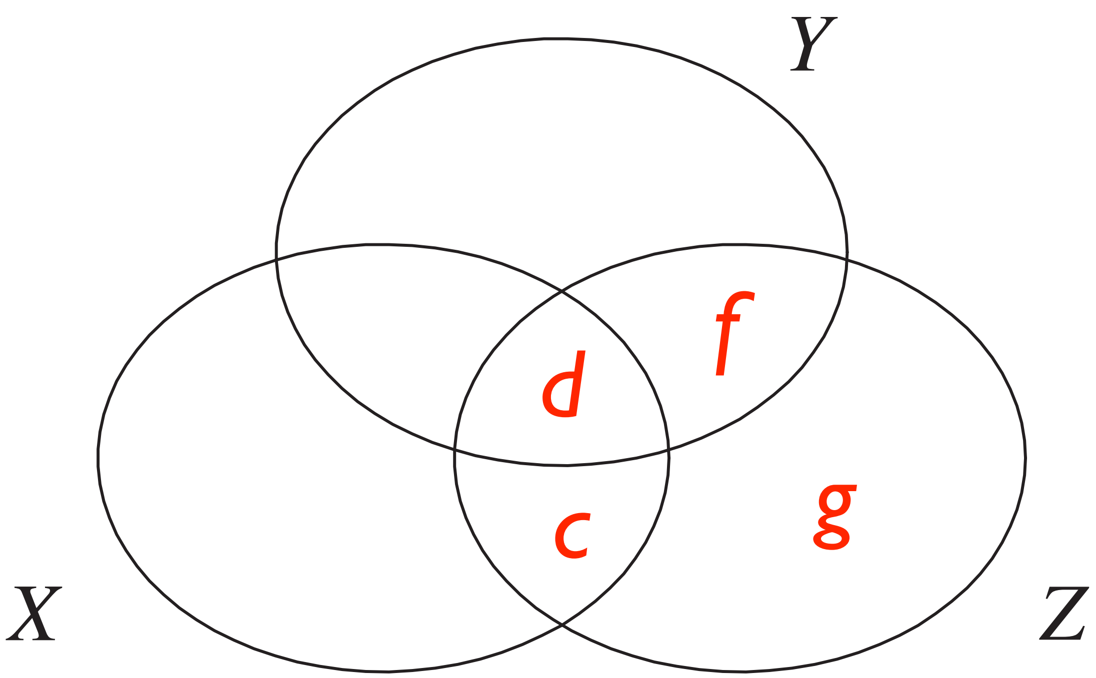


- Fill the nonempty atoms of the information diagram with arbitrary nonnegative numbers $a, b, \cdots, g$.
- Theorem 3.11 says that there exist r.v.'s X, Y, and Z whose I-measure μ^* is as shown.
- This can be seen by considering mutually independent r.v.'s $A, B, \cdots, G$ with entropies $a, b, \cdots, g$, respectively, and let

$$X = (A, B, C, D)$$
$$Y = (B, D, E, F)$$

Theorem 3.11 If there is no constraint on $X_1, X_2, \cdots, X_n$, then μ^* can take any set of nonnegative values on the nonempty atoms of $\mathcal{F}_n$.

Idea



- Fill the nonempty atoms of the information diagram with arbitrary nonnegative numbers $a, b, \cdots, g$.
- Theorem 3.11 says that there exist r.v.'s X, Y, and Z whose I-measure μ^* is as shown.
- This can be seen by considering mutually independent r.v.'s $A, B, \cdots, G$ with entropies $a, b, \cdots, g$, respectively, and let

$$
\begin{aligned}
X &= (A, B, C, D) \\
Y &= (B, D, E, F) \\
Z &= (C, D, F, G).
\end{aligned}
$$

Theorem 3.11 If there is no constraint on $X_1, X_2, \cdots, X_n$, then μ^* can take any set of nonnegative values on the nonempty atoms of $\mathcal{F}_n$.

Proof

Theorem 3.11 If there is no constraint on $X_1, X_2, \cdots, X_n$, then μ^* can take any set of nonnegative values on the nonempty atoms of $\mathcal{F}_n$.

Proof

1. Let $Y_A, A \in \mathcal{A}$ be mutually independent random variables. Note that these r.v.'s are labeled by the nonempty atoms of $\mathcal{F}_n$.

Theorem 3.11 If there is no constraint on $X_1, X_2, \cdots, X_n$, then μ^* can take any set of nonnegative values on the nonempty atoms of $\mathcal{F}_n$.

Proof

1. Let $Y_A, A \in \mathcal{A}$ be mutually independent random variables. Note that these r.v.'s are labeled by the nonempty atoms of $\mathcal{F}_n$.

2. Define the random variables $X_i, i = 1, 2, \cdots, n$ by

Theorem 3.11 If there is no constraint on $X_1, X_2, \cdots, X_n$, then μ^* can take any set of nonnegative values on the nonempty atoms of $\mathcal{F}_n$.

Proof

1. Let $Y_A, A \in \mathcal{A}$ be mutually independent random variables. Note that these r.v.'s are labeled by the nonempty atoms of $\mathcal{F}_n$.

2. Define the random variables $X_i, i = 1, 2, \cdots, n$ by

$$X_i = (Y_A : A \in \mathcal{A} \text{ and } A \subset \tilde{X}_i).$$

Theorem 3.11 If there is no constraint on $X_1, X_2, \cdots, X_n$, then μ^* can take any set of nonnegative values on the nonempty atoms of $\mathcal{F}_n$.

Proof

1. Let $Y_A, A \in \mathcal{A}$ be mutually independent random variables. Note that these r.v.'s are labeled by the nonempty atoms of $\mathcal{F}_n$.

2. Define the random variables $X_i, i = 1, 2, \cdots, n$ by

$$X_i = (Y_A : A \in \mathcal{A} \text{ and } A \subset \tilde{X}_i).$$

We will determine the I-Measure μ^* for $X_1, X_2, \cdots, X_n$ so defined.

Theorem 3.11 If there is no constraint on $X_1, X_2, \cdots, X_n$, then μ^* can take any set of nonnegative values on the nonempty atoms of $\mathcal{F}_n$.

Proof

1. Let $Y_A, A \in \mathcal{A}$ be mutually independent random variables. Note that these r.v.'s are labeled by the nonempty atoms of $\mathcal{F}_n$.

2. Define the random variables $X_i, i = 1, 2, \cdots, n$ by

$$X_i = (Y_A : A \in \mathcal{A} \text{ and } A \subset \tilde{X}_i).$$

We will determine the I-Measure μ^* for $X_1, X_2, \cdots, X_n$ so defined.

3. Since Y_A are mutually independent, for any nonempty subsets G of $\mathcal{N}_n$, we have

Theorem 3.11 If there is no constraint on $X_1, X_2, \cdots, X_n$, then μ^* can take any set of nonnegative values on the nonempty atoms of $\mathcal{F}_n$.

Proof

1. Let $Y_A, A \in \mathcal{A}$ be mutually independent random variables. Note that these r.v.'s are labeled by the nonempty atoms of $\mathcal{F}_n$.

2. Define the random variables $X_i, i = 1, 2, \cdots, n$ by

$$X_i = (Y_A : A \in \mathcal{A} \text{ and } A \subset \tilde{X}_i).$$

We will determine the I-Measure μ^* for $X_1, X_2, \cdots, X_n$ so defined.

3. Since Y_A are mutually independent, for any nonempty subsets G of $\mathcal{N}_n$, we have

$$H(X_G) \quad = \quad H(X_i, i \in G)$$

Theorem 3.11 If there is no constraint on $X_1, X_2, \cdots, X_n$, then μ^* can take any set of nonnegative values on the nonempty atoms of $\mathcal{F}_n$.

Proof

1. Let $Y_A, A \in \mathcal{A}$ be mutually independent random variables. Note that these r.v.'s are labeled by the nonempty atoms of $\mathcal{F}_n$.

2. Define the random variables $X_i, i = 1, 2, \cdots, n$ by

$$X_i = (Y_A : A \in \mathcal{A} \text{ and } A \subset \tilde{X}_i).$$

We will determine the I-Measure μ^* for $X_1, X_2, \cdots, X_n$ so defined.

3. Since Y_A are mutually independent, for any nonempty subsets G of $\mathcal{N}_n$, we have

$$H(X_G) \quad = \quad H(X_i, i \in G)$$

Theorem 3.11 If there is no constraint on $X_1, X_2, \cdots, X_n$, then μ^* can take any set of non-negative values on the nonempty atoms of $\mathcal{F}_n$.

Proof

1. Let $Y_A, A \in \mathcal{A}$ be mutually independent random variables. Note that these r.v.'s are labeled by the nonempty atoms of $\mathcal{F}_n$.

2. Define the random variables $X_i, i = 1, 2, \cdots, n$ by

$$X_i = (Y_A : A \in \mathcal{A} \text{ and } A \subset \tilde{X}_i).$$

We will determine the I-Measure μ^* for $X_1, X_2, \cdots, X_n$ so defined.

3. Since Y_A are mutually independent, for any nonempty subsets G of $\mathcal{N}_n$, we have

$$\begin{aligned} H(X_G) &= H(X_i, i \in G) \\ &= H((Y_A : A \in \mathcal{A} \text{ and } A \subset \tilde{X}_i), i \in G) \end{aligned}$$

Theorem 3.11 If there is no constraint on $X_1, X_2, \cdots, X_n$, then μ^* can take any set of nonnegative values on the nonempty atoms of $\mathcal{F}_n$.

Proof

1. Let $Y_A, A \in \mathcal{A}$ be mutually independent random variables. Note that these r.v.'s are labeled by the nonempty atoms of $\mathcal{F}_n$.

2. Define the random variables $X_i, i = 1, 2, \cdots, n$ by

$$X_i = (Y_A : A \in \mathcal{A} \text{ and } A \subset \tilde{X}_i).$$

We will determine the I-Measure μ^* for $X_1, X_2, \cdots,$ X_n so defined.

3. Since Y_A are mutually independent, for any nonempty subsets G of $\mathcal{N}_n$, we have

$$\begin{aligned} H(X_G) &= H(X_i, i \in G) \\ &= H((Y_A : A \in \mathcal{A} \text{ and } A \subset \tilde{X}_i), i \in G) \end{aligned}$$

Theorem 3.11 If there is no constraint on $X_1, X_2, \cdots, X_n$, then μ^* can take any set of non-negative values on the nonempty atoms of $\mathcal{F}_n$.

Proof

1. Let $Y_A, A \in \mathcal{A}$ be mutually independent random variables. Note that these r.v.'s are labeled by the nonempty atoms of $\mathcal{F}_n$.

2. Define the random variables $X_i, i = 1, 2, \cdots, n$ by

$$X_i = (Y_A : A \in \mathcal{A} \text{ and } A \subset \tilde{X}_i).$$

We will determine the I-Measure μ^* for $X_1, X_2, \cdots,$ X_n so defined.

3. Since Y_A are mutually independent, for any nonempty subsets G of $\mathcal{N}_n$, we have

$$\begin{aligned} H(X_G) &= H(X_i, i \in G) \\ &= H((Y_A : A \in \mathcal{A} \text{ and } \underline{A \subset \tilde{X}_i), i \in G}) \end{aligned}$$

Theorem 3.11 If there is no constraint on $X_1, X_2, \cdots, X_n$, then μ^* can take any set of nonnegative values on the nonempty atoms of $\mathcal{F}_n$.

Proof

1. Let $Y_A, A \in \mathcal{A}$ be mutually independent random variables. Note that these r.v.'s are labeled by the nonempty atoms of $\mathcal{F}_n$.

2. Define the random variables $X_i, i = 1, 2, \cdots, n$ by

$$X_i = (Y_A : A \in \mathcal{A} \text{ and } A \subset \tilde{X}_i).$$

We will determine the I-Measure μ^* for $X_1, X_2, \cdots, X_n$ so defined.

3. Since Y_A are mutually independent, for any nonempty subsets G of $\mathcal{N}_n$, we have

$$\begin{aligned} H(X_G) &= H(X_i, i \in G) \\ &= H((Y_A : A \in \mathcal{A} \text{ and } A \subset \tilde{X}_i), i \in G) \\ &= H(Y_A : A \in \mathcal{A} \text{ and } \underline{A \subset \tilde{X}_G}) \end{aligned}$$

Theorem 3.11 If there is no constraint on $X_1, X_2, \cdots, X_n$, then μ^* can take any set of nonnegative values on the nonempty atoms of $\mathcal{F}_n$.

Proof

1. Let $Y_A, A \in \mathcal{A}$ be mutually independent random variables. Note that these r.v.'s are labeled by the nonempty atoms of $\mathcal{F}_n$.

2. Define the random variables $X_i, i = 1, 2, \cdots, n$ by

$$X_i = (Y_A : A \in \mathcal{A} \text{ and } A \subset \tilde{X}_i).$$

We will determine the I-Measure μ^* for $X_1, X_2, \cdots, X_n$ so defined.

3. Since Y_A are mutually independent, for any nonempty subsets G of $\mathcal{N}_n$, we have

$$\begin{aligned} H(X_G) &= H(X_i, i \in G) \\ &= H((Y_A : A \in \mathcal{A} \text{ and } A \subset \tilde{X}_i), i \in G) \\ &= H(Y_A : A \in \mathcal{A} \text{ and } \underline{A \subset \tilde{X}_G}) \\ &= \sum_{A \in \mathcal{A} : \underline{A \subset \tilde{X}_G}} H(Y_A). \end{aligned} \tag{1}$$

Theorem 3.11 If there is no constraint on $X_1, X_2, \cdots, X_n$, then μ^* can take any set of nonnegative values on the nonempty atoms of $\mathcal{F}_n$.

Proof

1. Let $Y_A, A \in \mathcal{A}$ be mutually independent random variables. Note that these r.v.'s are labeled by the nonempty atoms of $\mathcal{F}_n$.

2. Define the random variables $X_i, i = 1, 2, \cdots, n$ by

$$X_i = (Y_A : A \in \mathcal{A} \text{ and } A \subset \tilde{X}_i).$$

We will determine the I-Measure μ^* for $X_1, X_2, \cdots, X_n$ so defined.

3. Since Y_A are mutually independent, for any nonempty subsets G of $\mathcal{N}_n$, we have

$$\begin{aligned} H(X_G) &= H(X_i, i \in G) \\ &= H((Y_A : A \in \mathcal{A} \text{ and } A \subset \tilde{X}_i), i \in G) \\ &= H(Y_A : A \in \mathcal{A} \text{ and } A \subset \tilde{X}_G) \\ &= \sum_{A \in \mathcal{A} : A \subset \tilde{X}_G} H(Y_A). \qquad (1) \end{aligned}$$

4. On the other hand, by set-additivity, we have

Theorem 3.11 If there is no constraint on $X_1, X_2, \cdots, X_n$, then μ^* can take any set of nonnegative values on the nonempty atoms of $\mathcal{F}_n$.

Proof

1. Let $Y_A, A \in \mathcal{A}$ be mutually independent random variables. Note that these r.v.'s are labeled by the nonempty atoms of $\mathcal{F}_n$.

2. Define the random variables $X_i, i = 1, 2, \cdots, n$ by

$$X_i = (Y_A : A \in \mathcal{A} \text{ and } A \subset \tilde{X}_i).$$

We will determine the I-Measure μ^* for $X_1, X_2, \cdots,$ X_n so defined.

3. Since Y_A are mutually independent, for any nonempty subsets G of $\mathcal{N}_n$, we have

$$\begin{aligned} H(X_G) &= H(X_i, i \in G) \\ &= H((Y_A : A \in \mathcal{A} \text{ and } A \subset \tilde{X}_i), i \in G) \\ &= H(Y_A : A \in \mathcal{A} \text{ and } A \subset \tilde{X}_G) \\ &= \sum_{A \in \mathcal{A} : A \subset \tilde{X}_G} H(Y_A). \end{aligned} \tag{1}$$

4. On the other hand, by set-additivity, we have

$$H(X_G) = \mu^*(\tilde{X}_G) = \sum_{A \in \mathcal{A} : A \subset \tilde{X}_G} \mu^*(A). \tag{2}$$

Theorem 3.11 If there is no constraint on $X_1, X_2, \cdots, X_n$, then μ^* can take any set of nonnegative values on the nonempty atoms of $\mathcal{F}_n$.

Proof

1. Let $Y_A, A \in \mathcal{A}$ be mutually independent random variables. Note that these r.v.'s are labeled by the nonempty atoms of $\mathcal{F}_n$.

2. Define the random variables $X_i, i = 1, 2, \cdots, n$ by

$$X_i = (Y_A : A \in \mathcal{A} \text{ and } A \subset \tilde{X}_i).$$

We will determine the I-Measure μ^* for $X_1, X_2, \cdots, X_n$ so defined.

3. Since Y_A are mutually independent, for any nonempty subsets G of $\mathcal{N}_n$, we have

$$\begin{aligned} H(X_G) &= H(X_i, i \in G) \\ &= H((Y_A : A \in \mathcal{A} \text{ and } A \subset \tilde{X}_i), i \in G) \\ &= H(Y_A : A \in \mathcal{A} \text{ and } A \subset \tilde{X}_G) \\ &= \sum_{A \in \mathcal{A} : A \subset \tilde{X}_G} H(Y_A). \qquad (1) \end{aligned}$$

4. On the other hand, by set-additivity, we have

$$H(X_G) = \mu^*(\tilde{X}_G) = \sum_{A \in \mathcal{A} : \underline{A \subset \tilde{X}_G}} \mu^*(A). \qquad (2)$$

Theorem 3.11 If there is no constraint on $X_1, X_2, \cdots, X_n$, then μ^* can take any set of nonnegative values on the nonempty atoms of $\mathcal{F}_n$.

Proof

1. Let $Y_A, A \in \mathcal{A}$ be mutually independent random variables. Note that these r.v.'s are labeled by the nonempty atoms of $\mathcal{F}_n$.

2. Define the random variables $X_i, i = 1, 2, \cdots, n$ by

$$X_i = (Y_A : A \in \mathcal{A} \text{ and } A \subset \tilde{X}_i).$$

We will determine the I-Measure μ^* for $X_1, X_2, \cdots, X_n$ so defined.

3. Since Y_A are mutually independent, for any nonempty subsets G of $\mathcal{N}_n$, we have

$$\begin{aligned} H(X_G) &= H(X_i, i \in G) \\ &= H((Y_A : A \in \mathcal{A} \text{ and } A \subset \tilde{X}_i), i \in G) \\ &= H(Y_A : A \in \mathcal{A} \text{ and } A \subset \tilde{X}_G) \\ &= \sum_{A \in \mathcal{A} : A \subset \tilde{X}_G} H(Y_A). \end{aligned} \tag{1}$$

4. On the other hand, by set-additivity, we have

$$H(X_G) = \mu^*(\tilde{X}_G) = \sum_{A \in \mathcal{A} : A \subset \tilde{X}_G} \mu^*(A). \tag{2}$$

Equating the right hand sides of (1) and (2), we have

Theorem 3.11 If there is no constraint on $X_1, X_2, \cdots, X_n$, then μ^* can take any set of nonnegative values on the nonempty atoms of $\mathcal{F}_n$.

Proof

1. Let $Y_A, A \in \mathcal{A}$ be mutually independent random variables. Note that these r.v.'s are labeled by the nonempty atoms of $\mathcal{F}_n$.

2. Define the random variables $X_i, i = 1, 2, \cdots, n$ by

$$X_i = (Y_A : A \in \mathcal{A} \text{ and } A \subset \tilde{X}_i).$$

We will determine the I-Measure μ^* for $X_1, X_2, \cdots, X_n$ so defined.

3. Since Y_A are mutually independent, for any nonempty subsets G of $\mathcal{N}_n$, we have

$$\begin{aligned} H(X_G) &= H(X_i, i \in G) \\ &= H((Y_A : A \in \mathcal{A} \text{ and } A \subset \tilde{X}_i), i \in G) \\ &= H(Y_A : A \in \mathcal{A} \text{ and } A \subset \tilde{X}_G) \\ &= \sum_{A \in \mathcal{A} : A \subset \tilde{X}_G} H(Y_A). \qquad (1) \end{aligned}$$

4. On the other hand, by set-additivity, we have

$$H(X_G) = \mu^*(\tilde{X}_G) = \sum_{A \in \mathcal{A} : A \subset \tilde{X}_G} \mu^*(A). \qquad (2)$$

Equating the right hand sides of (1) and (2), we have

Theorem 3.11 If there is no constraint on $X_1, X_2, \cdots, X_n$, then μ^* can take any set of non-negative values on the nonempty atoms of $\mathcal{F}_n$.

Proof

1. Let $Y_A, A \in \mathcal{A}$ be mutually independent random variables. Note that these r.v.'s are labeled by the nonempty atoms of $\mathcal{F}_n$.

2. Define the random variables $X_i, i = 1, 2, \cdots, n$ by

$$X_i = (Y_A : A \in \mathcal{A} \text{ and } A \subset \tilde{X}_i).$$

We will determine the I-Measure μ^* for $X_1, X_2, \cdots, X_n$ so defined.

3. Since Y_A are mutually independent, for any nonempty subsets G of $\mathcal{N}_n$, we have

$$\begin{aligned} H(X_G) &= H(X_i, i \in G) \\ &= H((Y_A : A \in \mathcal{A} \text{ and } A \subset \tilde{X}_i), i \in G) \\ &= H(Y_A : A \in \mathcal{A} \text{ and } A \subset \tilde{X}_G) \\ &= \underline{\sum_{A \in \mathcal{A} : A \subset \tilde{X}_G} H(Y_A)}. \end{aligned} \quad (1)$$

4. On the other hand, by set-additivity, we have

$$H(X_G) = \mu^*(\tilde{X}_G) = \underline{\sum_{A \in \mathcal{A} : A \subset \tilde{X}_G} \mu^*(A)}. \quad (2)$$

Equating the right hand sides of (1) and (2), we have

Theorem 3.11 If there is no constraint on $X_1, X_2, \cdots, X_n$, then μ^* can take any set of nonnegative values on the nonempty atoms of $\mathcal{F}_n$.

Proof

1. Let $Y_A, A \in \mathcal{A}$ be mutually independent random variables. Note that these r.v.'s are labeled by the nonempty atoms of $\mathcal{F}_n$.

2. Define the random variables $X_i, i = 1, 2, \cdots, n$ by

$$X_i = (Y_A : A \in \mathcal{A} \text{ and } A \subset \tilde{X}_i).$$

We will determine the I-Measure μ^* for $X_1, X_2, \cdots, X_n$ so defined.

3. Since Y_A are mutually independent, for any nonempty subsets G of $\mathcal{N}_n$, we have

$$\begin{aligned} H(X_G) &= H(X_i, i \in G) \\ &= H((Y_A : A \in \mathcal{A} \text{ and } A \subset \tilde{X}_i), i \in G) \\ &= H(Y_A : A \in \mathcal{A} \text{ and } A \subset \tilde{X}_G) \\ &= \sum_{A \in \mathcal{A} : A \subset \tilde{X}_G} H(Y_A). \end{aligned} \qquad (1)$$

4. On the other hand, by set-additivity, we have

$$H(X_G) = \mu^*(\tilde{X}_G) = \sum_{A \in \mathcal{A} : A \subset \tilde{X}_G} \mu^*(A). \qquad (2)$$

Equating the right hand sides of (1) and (2), we have

$$\sum_{A \in \mathcal{A} : A \subset \tilde{X}_G} H(Y_A) = \sum_{A \in \mathcal{A} : A \subset \tilde{X}_G} \mu^*(A).$$

Theorem 3.11 If there is no constraint on $X_1, X_2, \cdots, X_n$, then μ^* can take any set of non-negative values on the nonempty atoms of $\mathcal{F}_n$.

Proof

1. Let $Y_A, A \in \mathcal{A}$ be mutually independent random variables. Note that these r.v.'s are labeled by the nonempty atoms of $\mathcal{F}_n$.

2. Define the random variables $X_i, i = 1, 2, \cdots, n$ by

$$X_i = (Y_A : A \in \mathcal{A} \text{ and } A \subset \tilde{X}_i).$$

We will determine the I-Measure μ^* for $X_1, X_2, \cdots, X_n$ so defined.

3. Since Y_A are mutually independent, for any nonempty subsets G of $\mathcal{N}_n$, we have

$$\begin{aligned} H(X_G) &= H(X_i, i \in G) \\ &= H((Y_A : A \in \mathcal{A} \text{ and } A \subset \tilde{X}_i), i \in G) \\ &= H(Y_A : A \in \mathcal{A} \text{ and } A \subset \tilde{X}_G) \\ &= \sum_{A \in \mathcal{A} : A \subset \tilde{X}_G} H(Y_A). \qquad (1) \end{aligned}$$

4. On the other hand, by set-additivity, we have

$$H(X_G) = \mu^*(\tilde{X}_G) = \sum_{A \in \mathcal{A} : A \subset \tilde{X}_G} \mu^*(A). \qquad (2)$$

Equating the right hand sides of (1) and (2), we have

$$\sum_{A \in \mathcal{A} : A \subset \tilde{X}_G} H(Y_A) = \sum_{A \in \mathcal{A} : A \subset \tilde{X}_G} \mu^*(A).$$

Theorem 3.11 If there is no constraint on $X_1, X_2, \cdots, X_n$, then μ^* can take any set of nonnegative values on the nonempty atoms of $\mathcal{F}_n$.

Proof

1. Let $Y_A, A \in \mathcal{A}$ be mutually independent random variables. Note that these r.v.'s are labeled by the nonempty atoms of $\mathcal{F}_n$.

2. Define the random variables $X_i, i = 1, 2, \cdots, n$ by

$$X_i = (Y_A : A \in \mathcal{A} \text{ and } A \subset \tilde{X}_i).$$

We will determine the I-Measure μ^* for $X_1, X_2, \cdots, X_n$ so defined.

3. Since Y_A are mutually independent, for any nonempty subsets G of $\mathcal{N}_n$, we have

$$\begin{aligned} H(X_G) &= H(X_i, i \in G) \\ &= H((Y_A : A \in \mathcal{A} \text{ and } A \subset \tilde{X}_i), i \in G) \\ &= H(Y_A : A \in \mathcal{A} \text{ and } A \subset \tilde{X}_G) \\ &= \sum_{A \in \mathcal{A} : A \subset \tilde{X}_G} H(Y_A). \qquad (1) \end{aligned}$$

4. On the other hand, by set-additivity, we have

$$H(X_G) = \mu^*(\tilde{X}_G) = \sum_{A \in \mathcal{A} : A \subset \tilde{X}_G} \mu^*(A). \qquad (2)$$

Equating the right hand sides of (1) and (2), we have

$$\sum_{A \in \mathcal{A} : \underline{A \subset \tilde{X}_G}} H(Y_A) = \sum_{A \in \mathcal{A} : \underline{A \subset \tilde{X}_G}} \mu^*(A).$$

Theorem 3.11 If there is no constraint on $X_1, X_2, \cdots, X_n$, then μ^* can take any set of nonnegative values on the nonempty atoms of $\mathcal{F}_n$.

Proof

1. Let $Y_A, A \in \mathcal{A}$ be mutually independent random variables. Note that these r.v.'s are labeled by the nonempty atoms of $\mathcal{F}_n$.

2. Define the random variables $X_i, i = 1, 2, \cdots, n$ by

$$X_i = (Y_A : A \in \mathcal{A} \text{ and } A \subset \tilde{X}_i).$$

We will determine the I-Measure μ^* for $X_1, X_2, \cdots, X_n$ so defined.

3. Since Y_A are mutually independent, for any nonempty subsets G of $\mathcal{N}_n$, we have

$$\begin{aligned} H(X_G) &= H(X_i, i \in G) \\ &= H((Y_A : A \in \mathcal{A} \text{ and } A \subset \tilde{X}_i), i \in G) \\ &= H(Y_A : A \in \mathcal{A} \text{ and } A \subset \tilde{X}_G) \\ &= \sum_{A \in \mathcal{A} : A \subset \tilde{X}_G} H(Y_A). \end{aligned} \tag{1}$$

4. On the other hand, by set-additivity, we have

$$H(X_G) = \mu^*(\tilde{X}_G) = \sum_{A \in \mathcal{A} : A \subset \tilde{X}_G} \mu^*(A). \tag{2}$$

Equating the right hand sides of (1) and (2), we have

$$\sum_{A \in \mathcal{A} : A \subset \tilde{X}_G} H(Y_A) = \sum_{A \in \mathcal{A} : A \subset \tilde{X}_G} \mu^*(A).$$

5. Here $\mu^*(A)$ are the unknowns. Evidently, we can make the above equality hold for all nonempty subsets G of $\mathcal{N}_n$ by taking

Theorem 3.11 If there is no constraint on $X_1, X_2, \cdots, X_n$, then μ^* can take any set of nonnegative values on the nonempty atoms of $\mathcal{F}_n$.

Proof

1. Let $Y_A, A \in \mathcal{A}$ be mutually independent random variables. Note that these r.v.'s are labeled by the nonempty atoms of $\mathcal{F}_n$.

2. Define the random variables $X_i, i = 1, 2, \cdots, n$ by

$$X_i = (Y_A : A \in \mathcal{A} \text{ and } A \subset \tilde{X}_i).$$

We will determine the I-Measure μ^* for $X_1, X_2, \cdots, X_n$ so defined.

3. Since Y_A are mutually independent, for any nonempty subsets G of $\mathcal{N}_n$, we have

$$\begin{aligned} H(X_G) &= H(X_i, i \in G) \\ &= H((Y_A : A \in \mathcal{A} \text{ and } A \subset \tilde{X}_i), i \in G) \\ &= H(Y_A : A \in \mathcal{A} \text{ and } A \subset \tilde{X}_G) \\ &= \sum_{A \in \mathcal{A} : A \subset \tilde{X}_G} H(Y_A). \qquad (1) \end{aligned}$$

4. On the other hand, by set-additivity, we have

$$H(X_G) = \mu^*(\tilde{X}_G) = \sum_{A \in \mathcal{A} : A \subset \tilde{X}_G} \mu^*(A). \qquad (2)$$

Equating the right hand sides of (1) and (2), we have

$$\sum_{A \in \mathcal{A} : A \subset \tilde{X}_G} H(Y_A) = \sum_{A \in \mathcal{A} : A \subset \tilde{X}_G} \mu^*(A).$$

5. Here $\mu^*(A)$ are the unknowns. Evidently, we can make the above equality hold for all nonempty subsets G of $\mathcal{N}_n$ by taking

$$\mu^*(A) = H(Y_A)$$

Theorem 3.11 If there is no constraint on $X_1, X_2, \cdots, X_n$, then μ^* can take any set of nonnegative values on the nonempty atoms of $\mathcal{F}_n$.

Proof

1. Let $Y_A, A \in \mathcal{A}$ be mutually independent random variables. Note that these r.v.'s are labeled by the nonempty atoms of $\mathcal{F}_n$.

2. Define the random variables $X_i, i = 1, 2, \cdots, n$ by

$$X_i = (Y_A : A \in \mathcal{A} \text{ and } A \subset \tilde{X}_i).$$

We will determine the I-Measure μ^* for $X_1, X_2, \cdots, X_n$ so defined.

3. Since Y_A are mutually independent, for any nonempty subsets G of $\mathcal{N}_n$, we have

$$\begin{aligned} H(X_G) &= H(X_i, i \in G) \\ &= H((Y_A : A \in \mathcal{A} \text{ and } A \subset \tilde{X}_i), i \in G) \\ &= H(Y_A : A \in \mathcal{A} \text{ and } A \subset \tilde{X}_G) \\ &= \sum_{A \in \mathcal{A} : A \subset \tilde{X}_G} H(Y_A). \qquad (1) \end{aligned}$$

4. On the other hand, by set-additivity, we have

$$H(X_G) = \mu^*(\tilde{X}_G) = \sum_{A \in \mathcal{A} : A \subset \tilde{X}_G} \mu^*(A). \qquad (2)$$

Equating the right hand sides of (1) and (2), we have

$$\sum_{A \in \mathcal{A} : A \subset \tilde{X}_G} H(Y_A) = \sum_{A \in \mathcal{A} : A \subset \tilde{X}_G} \mu^*(A).$$

5. Here $\mu^*(A)$ are the unknowns. Evidently, we can make the above equality hold for all nonempty subsets G of $\mathcal{N}_n$ by taking

$$\mu^*(A) = H(Y_A)$$

Theorem 3.11 If there is no constraint on $X_1, X_2, \cdots, X_n$, then μ^* can take any set of nonnegative values on the nonempty atoms of $\mathcal{F}_n$.

Proof

1. Let $Y_A, A \in \mathcal{A}$ be mutually independent random variables. Note that these r.v.'s are labeled by the nonempty atoms of $\mathcal{F}_n$.

2. Define the random variables $X_i, i = 1, 2, \cdots, n$ by

$$X_i = (Y_A : A \in \mathcal{A} \text{ and } A \subset \tilde{X}_i).$$

We will determine the I-Measure μ^* for $X_1, X_2, \cdots, X_n$ so defined.

3. Since Y_A are mutually independent, for any nonempty subsets G of $\mathcal{N}_n$, we have

$$\begin{aligned} H(X_G) &= H(X_i, i \in G) \\ &= H((Y_A : A \in \mathcal{A} \text{ and } A \subset \tilde{X}_i), i \in G) \\ &= H(Y_A : A \in \mathcal{A} \text{ and } A \subset \tilde{X}_G) \\ &= \sum_{A \in \mathcal{A} : A \subset \tilde{X}_G} H(Y_A). \end{aligned} \tag{1}$$

4. On the other hand, by set-additivity, we have

$$H(X_G) = \mu^*(\tilde{X}_G) = \sum_{A \in \mathcal{A} : A \subset \tilde{X}_G} \mu^*(A). \tag{2}$$

Equating the right hand sides of (1) and (2), we have

$$\sum_{A \in \mathcal{A} : A \subset \tilde{X}_G} H(Y_A) = \sum_{A \in \mathcal{A} : A \subset \tilde{X}_G} \mu^*(A).$$

5. Here $\mu^*(A)$ are the unknowns. Evidently, we can make the above equality hold for all nonempty subsets G of $\mathcal{N}_n$ by taking

$$\mu^*(A) = H(Y_A)$$

for all $A \in \mathcal{A}$. By the uniqueness of μ^* (Theorem 3.9), this is also the only possibility for μ^*.

Theorem 3.11 If there is no constraint on $X_1, X_2, \cdots, X_n$, then μ^* can take any set of nonnegative values on the nonempty atoms of $\mathcal{F}_n$.

Proof

1. Let $Y_A, A \in \mathcal{A}$ be mutually independent random variables. Note that these r.v.'s are labeled by the nonempty atoms of $\mathcal{F}_n$.

2. Define the random variables $X_i, i = 1, 2, \cdots, n$ by

$$X_i = (Y_A : A \in \mathcal{A} \text{ and } A \subset \tilde{X}_i).$$

We will determine the I-Measure μ^* for $X_1, X_2, \cdots, X_n$ so defined.

3. Since Y_A are mutually independent, for any nonempty subsets G of $\mathcal{N}_n$, we have

$$\begin{aligned} H(X_G) &= H(X_i, i \in G) \\ &= H((Y_A : A \in \mathcal{A} \text{ and } A \subset \tilde{X}_i), i \in G) \\ &= H(Y_A : A \in \mathcal{A} \text{ and } A \subset \tilde{X}_G) \\ &= \sum_{A \in \mathcal{A} : A \subset \tilde{X}_G} H(Y_A). \qquad (1) \end{aligned}$$

4. On the other hand, by set-additivity, we have

$$H(X_G) = \mu^*(\tilde{X}_G) = \sum_{A \in \mathcal{A} : A \subset \tilde{X}_G} \mu^*(A). \qquad (2)$$

Equating the right hand sides of (1) and (2), we have

$$\sum_{A \in \mathcal{A} : A \subset \tilde{X}_G} H(Y_A) = \sum_{A \in \mathcal{A} : A \subset \tilde{X}_G} \mu^*(A).$$

5. Here $\mu^*(A)$ are the unknowns. Evidently, we can make the above equality hold for all nonempty subsets G of $\mathcal{N}_n$ by taking

$$\mu^*(A) = H(Y_A)$$

for all $A \in \mathcal{A}$. By the uniqueness of μ^* (Theorem 3.9), this is also the only possibility for μ^*.

Theorem 3.9 μ^* is the unique signed measure on $\mathcal{F}_n$ which is consistent with all Shannon's information measures.

Theorem 3.11 If there is no constraint on $X_1, X_2, \cdots, X_n$, then μ^* can take any set of nonnegative values on the nonempty atoms of $\mathcal{F}_n$.

Proof

1. Let $Y_A, A \in \mathcal{A}$ be mutually independent random variables. Note that these r.v.'s are labeled by the nonempty atoms of $\mathcal{F}_n$.

2. Define the random variables $X_i, i = 1, 2, \cdots, n$ by

$$X_i = (Y_A : A \in \mathcal{A} \text{ and } A \subset \tilde{X}_i).$$

We will determine the I-Measure μ^* for $X_1, X_2, \cdots, X_n$ so defined.

3. Since Y_A are mutually independent, for any nonempty subsets G of $\mathcal{N}_n$, we have

$$\begin{aligned} H(X_G) &= H(X_i, i \in G) \\ &= H((Y_A : A \in \mathcal{A} \text{ and } A \subset \tilde{X}_i), i \in G) \\ &= H(Y_A : A \in \mathcal{A} \text{ and } A \subset \tilde{X}_G) \\ &= \sum_{A \in \mathcal{A} : A \subset \tilde{X}_G} H(Y_A). \end{aligned} \qquad (1)$$

4. On the other hand, by set-additivity, we have

$$H(X_G) = \mu^*(\tilde{X}_G) = \sum_{A \in \mathcal{A} : A \subset \tilde{X}_G} \mu^*(A). \qquad (2)$$

Equating the right hand sides of (1) and (2), we have

$$\sum_{A \in \mathcal{A} : A \subset \tilde{X}_G} H(Y_A) = \sum_{A \in \mathcal{A} : A \subset \tilde{X}_G} \mu^*(A).$$

5. Here $\mu^*(A)$ are the unknowns. Evidently, we can make the above equality hold for all nonempty subsets G of $\mathcal{N}_n$ by taking

$$\mu^*(A) = H(Y_A)$$

for all $A \in \mathcal{A}$. By the uniqueness of μ^* (Theorem 3.9), this is also the only possibility for μ^*.

6. Since $H(Y_A)$ can take any nonnegative value by Corollary 2.44, μ^* can take any set of nonnegative values on the nonempty atoms of $\mathcal{F}_n$. The theorem is proved.

Theorem 3.9 μ^* is the unique signed measure on $\mathcal{F}_n$ which is consistent with all Shannon's information measures.

Theorem 3.11 If there is no constraint on $X_1, X_2, \cdots, X_n$, then μ^* can take any set of nonnegative values on the nonempty atoms of $\mathcal{F}_n$.

Proof

1. Let $Y_A, A \in \mathcal{A}$ be mutually independent random variables. Note that these r.v.'s are labeled by the nonempty atoms of $\mathcal{F}_n$.

2. Define the random variables $X_i, i = 1, 2, \cdots, n$ by

$$X_i = (Y_A : A \in \mathcal{A} \text{ and } A \subset \tilde{X}_i).$$

We will determine the I-Measure μ^* for $X_1, X_2, \cdots, X_n$ so defined.

3. Since Y_A are mutually independent, for any nonempty subsets G of $\mathcal{N}_n$, we have

$$\begin{aligned} H(X_G) &= H(X_i, i \in G) \\ &= H((Y_A : A \in \mathcal{A} \text{ and } A \subset \tilde{X}_i), i \in G) \\ &= H(Y_A : A \in \mathcal{A} \text{ and } A \subset \tilde{X}_G) \\ &= \sum_{A \in \mathcal{A} : A \subset \tilde{X}_G} H(Y_A). \qquad (1) \end{aligned}$$

4. On the other hand, by set-additivity, we have

$$H(X_G) = \mu^*(\tilde{X}_G) = \sum_{A \in \mathcal{A} : A \subset \tilde{X}_G} \mu^*(A). \qquad (2)$$

Equating the right hand sides of (1) and (2), we have

$$\sum_{A \in \mathcal{A} : A \subset \tilde{X}_G} H(Y_A) = \sum_{A \in \mathcal{A} : A \subset \tilde{X}_G} \mu^*(A).$$

5. Here $\mu^*(A)$ are the unknowns. Evidently, we can make the above equality hold for all nonempty subsets G of $\mathcal{N}_n$ by taking

$$\mu^*(A) = H(Y_A)$$

for all $A \in \mathcal{A}$. By the uniqueness of μ^* (Theorem 3.9), this is also the only possibility for μ^*.

6. Since $H(Y_A)$ can take any nonnegative value by Corollary 2.44, μ^* can take any set of nonnegative values on the nonempty atoms of $\mathcal{F}_n$. The theorem is proved.

Theorem 3.9 μ^* is the unique signed measure on $\mathcal{F}_n$ which is consistent with all Shannon's information measures.

Corollary 2.44 The entropy of a random variable may take any nonnegative real value.

Information Diagrams for Markov Chains

Information Diagrams for Markov Chains

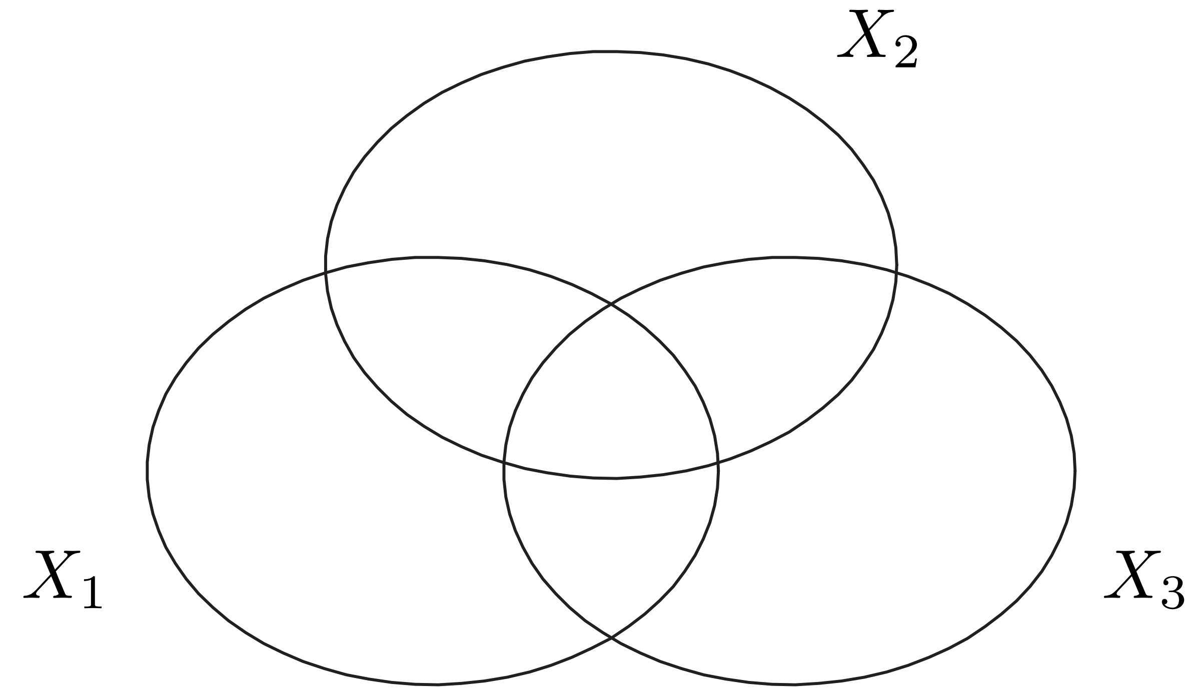
- If $X_1 \to X_2 \to \cdots \to X_n$ form a Markov chain, then the structure of μ^* is much simpler and hence the information diagram can be simplified.

Information Diagrams for Markov Chains

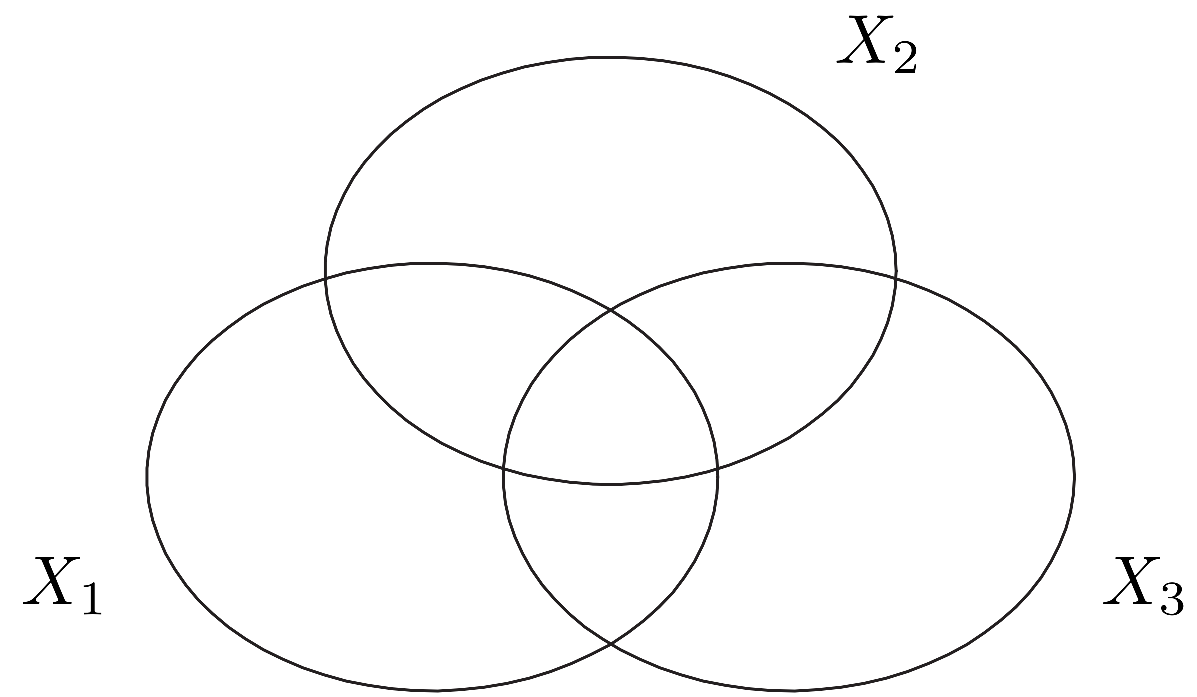
- If $X_1 \to X_2 \to \cdots \to X_n$ form a Markov chain, then the structure of μ^* is much simpler and hence the information diagram can be simplified.
- For $n = 3$, $X_1 \to X_2 \to X_3$ iff $I(X_1; X_3|X_2) = 0$, or $\mu^*(\tilde{X}_1 \cap \tilde{X}_3 - \tilde{X}_2) = 0$.

Information Diagrams for Markov Chains

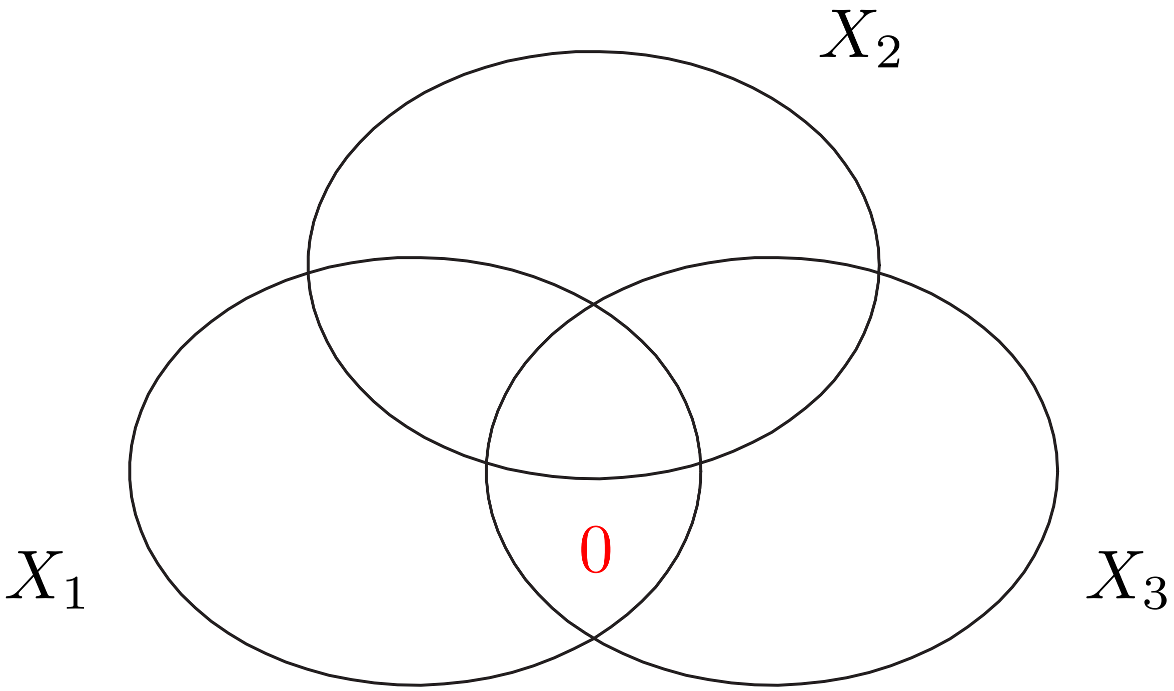
- If $X_1 \to X_2 \to \cdots \to X_n$ form a Markov chain, then the structure of μ^* is much simpler and hence the information diagram can be simplified.
- For $n = 3$, $X_1 \to X_2 \to X_3$ iff $I(X_1; X_3|X_2) = 0$, or $\mu^*(\tilde{X}_1 \cap \tilde{X}_3 - \tilde{X}_2) = 0$.
- So the atom $\tilde{X}_1 \cap \tilde{X}_3 - \tilde{X}_2$ can be suppressed in the information diagram.


X_2
X_1
X_3

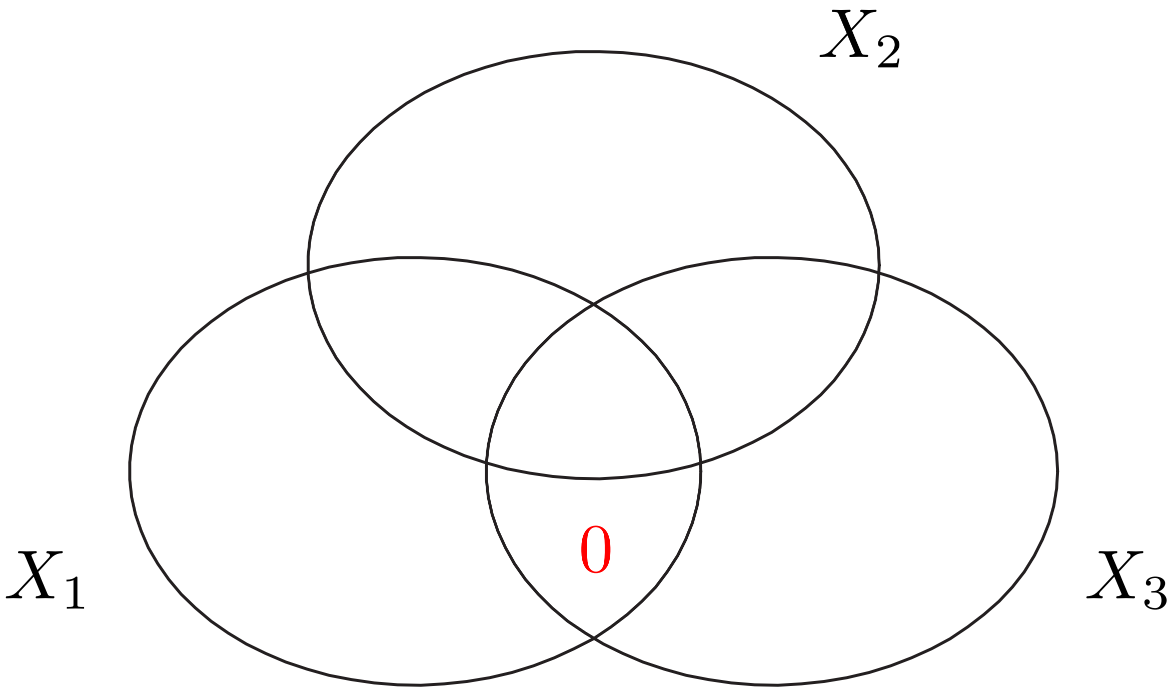
- Suppose $I(X_1; X_3|X_2) = 0$.



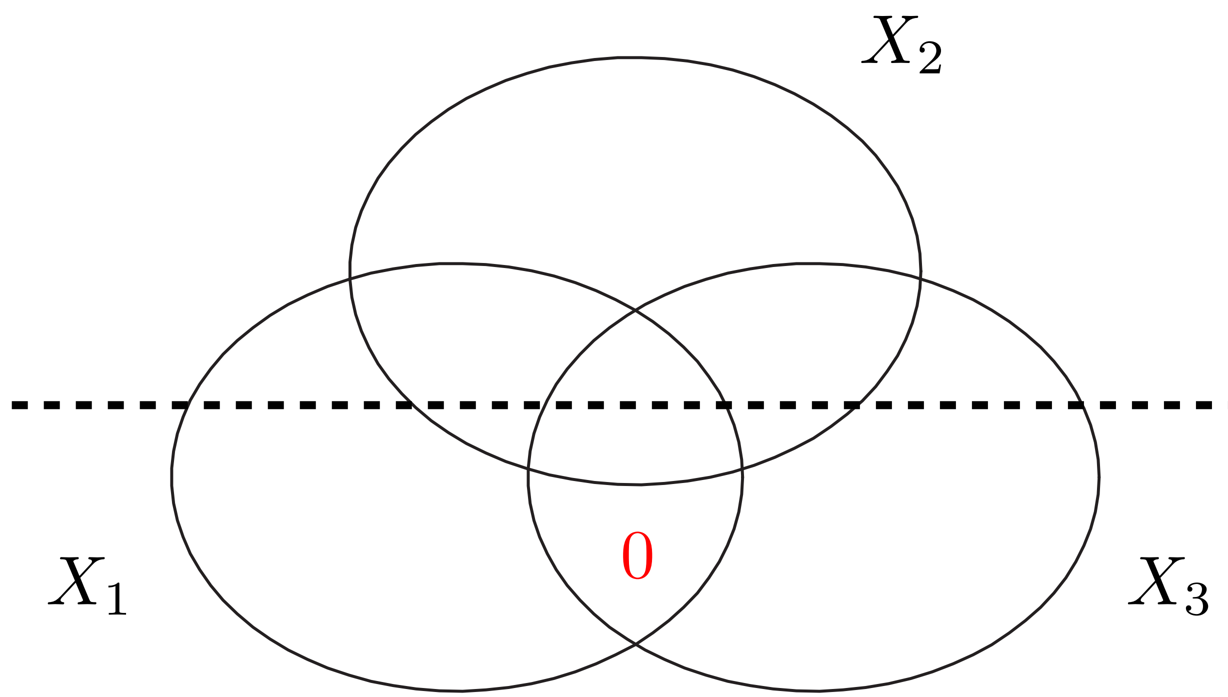
- Suppose $I(X_1; X_3|X_2) = 0$.
- Then $\mu^*(\tilde{X}_1 \cap \tilde{X}_3 - \tilde{X}_2) = 0$.



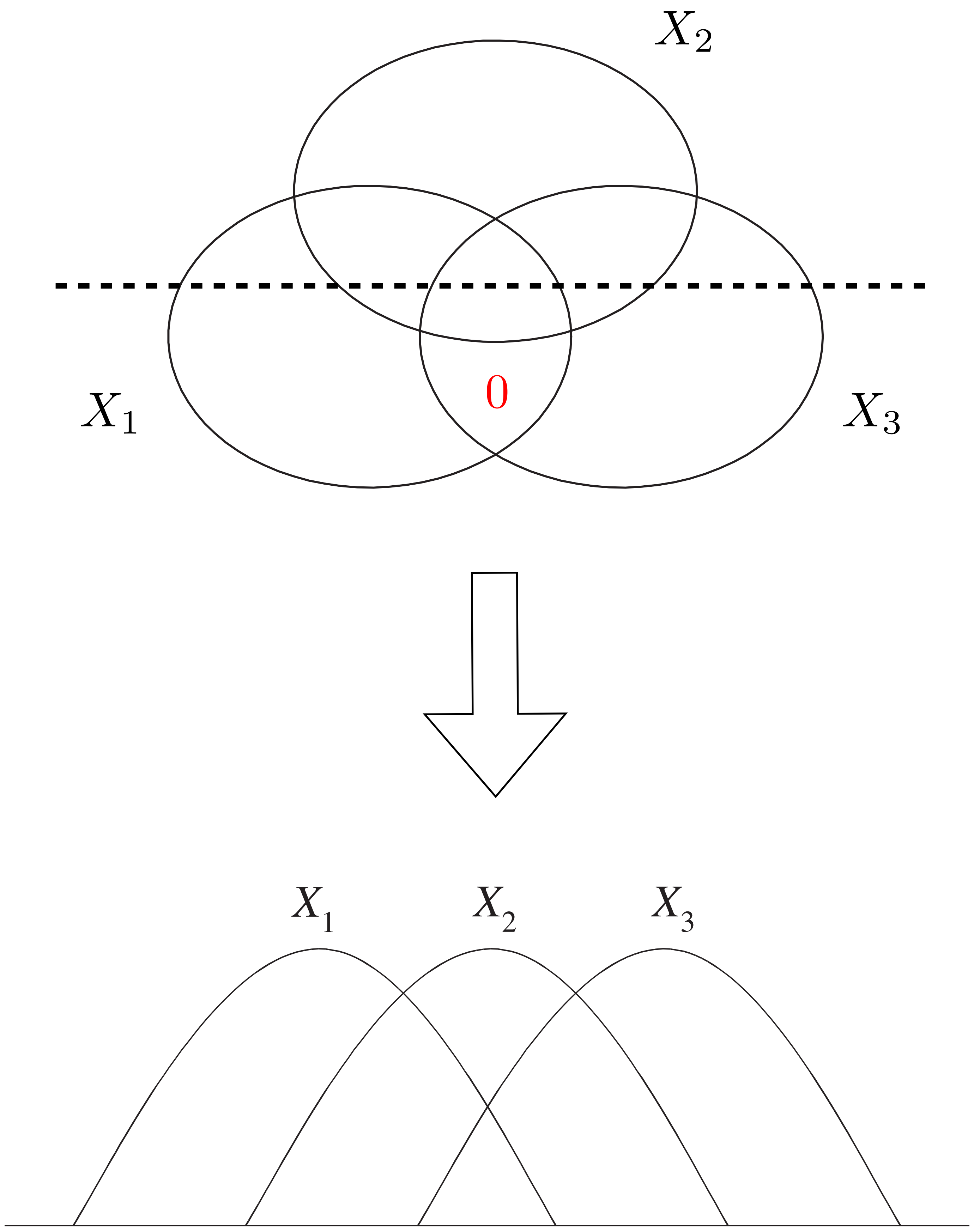
- Suppose $I(X_1; X_3|X_2) = 0$.
- Then $\mu^*(\tilde{X}_1 \cap \tilde{X}_3 - \tilde{X}_2) = 0$.
- Suppress the atom $\tilde{X}_1 \cap \tilde{X}_3 - \tilde{X}_2$ by setting it to $\emptyset$.



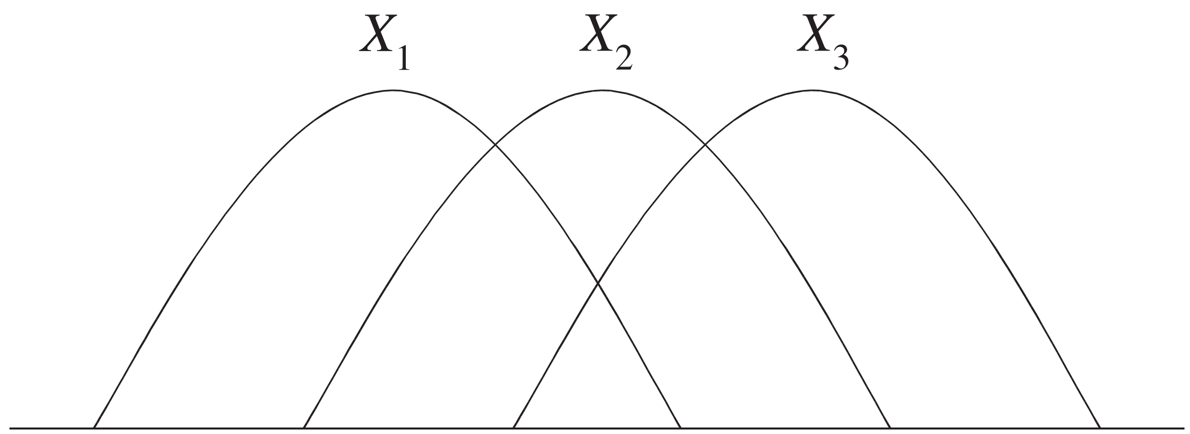
- Suppose $I(X_1; X_3|X_2) = 0$.
- Then $\mu^*(\tilde{X}_1 \cap \tilde{X}_3 - \tilde{X}_2) = 0$.
- Suppress the atom $\tilde{X}_1 \cap \tilde{X}_3 - \tilde{X}_2$ by setting it to $\emptyset$.



- Suppose $I(X_1; X_3|X_2) = 0$.
- Then $\mu^*(\tilde{X}_1 \cap \tilde{X}_3 - \tilde{X}_2) = 0$.
- Suppress the atom $\tilde{X}_1 \cap \tilde{X}_3 - \tilde{X}_2$ by setting it to $\emptyset$.

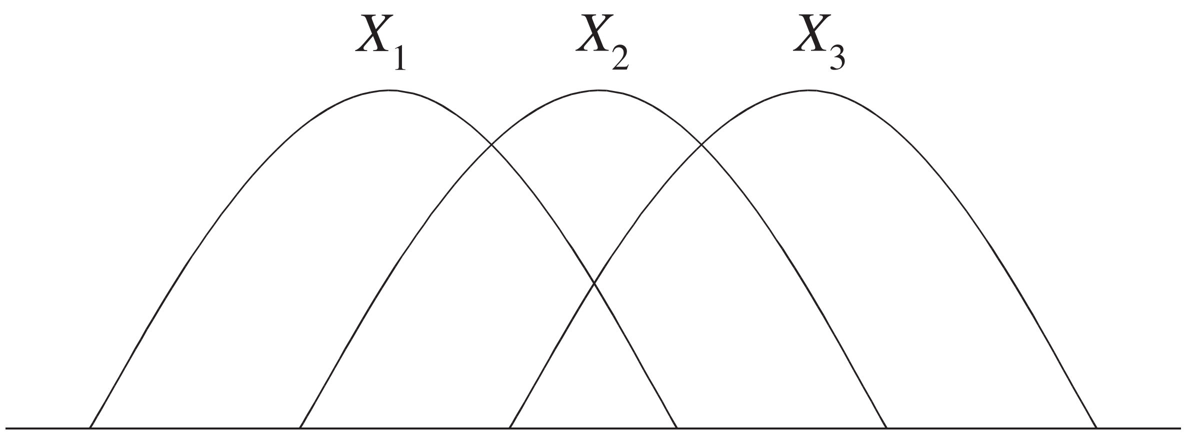


μ^* **for** $X_1 \rightarrow X_2 \rightarrow X_3$



μ^* for $X_1 \to X_2 \to X_3$

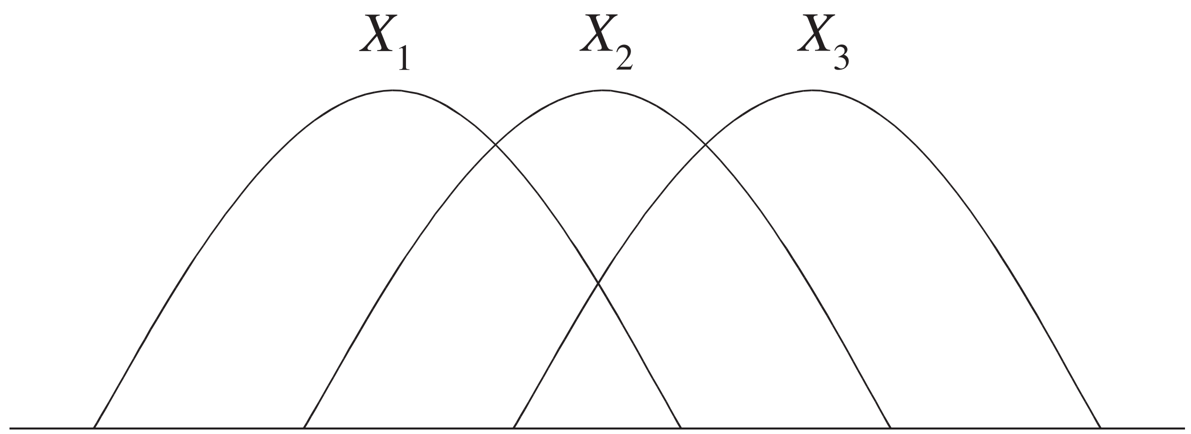
1. In this information diagram,



μ^* for $X_1 \to X_2 \to X_3$

1. In this information diagram,

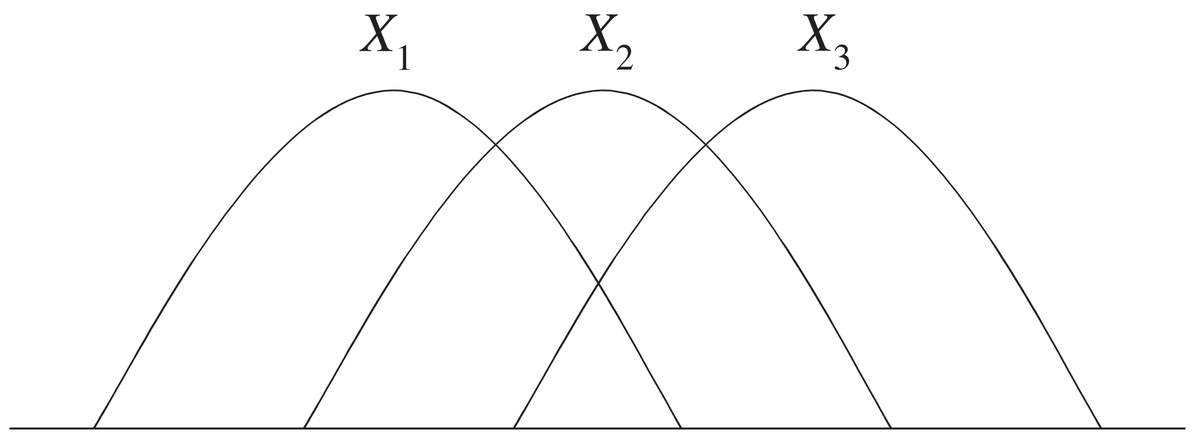
$$I(X_1;X_3|X_2) = \mu^*(\tilde{X}_1 \cap \tilde{X}_3 - \tilde{X}_2)$$



μ^* for $X_1 \rightarrow X_2 \rightarrow X_3$

1. In this information diagram,

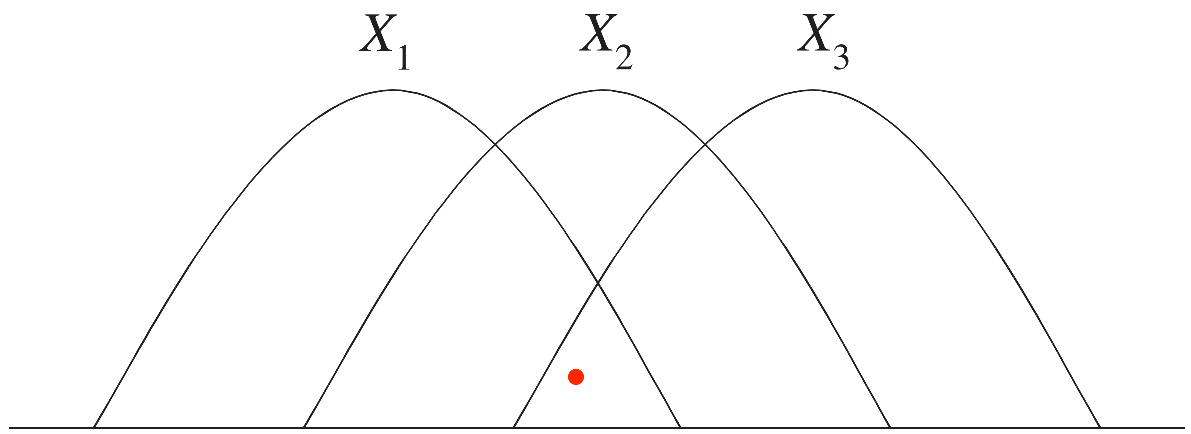
$$I(X_1; X_3|X_2) \quad = \quad \mu^*(\tilde{X}_1 \cap \tilde{X}_3 - \tilde{X}_2)$$



μ^* for $X_1 \rightarrow X_2 \rightarrow X_3$

1. In this information diagram,

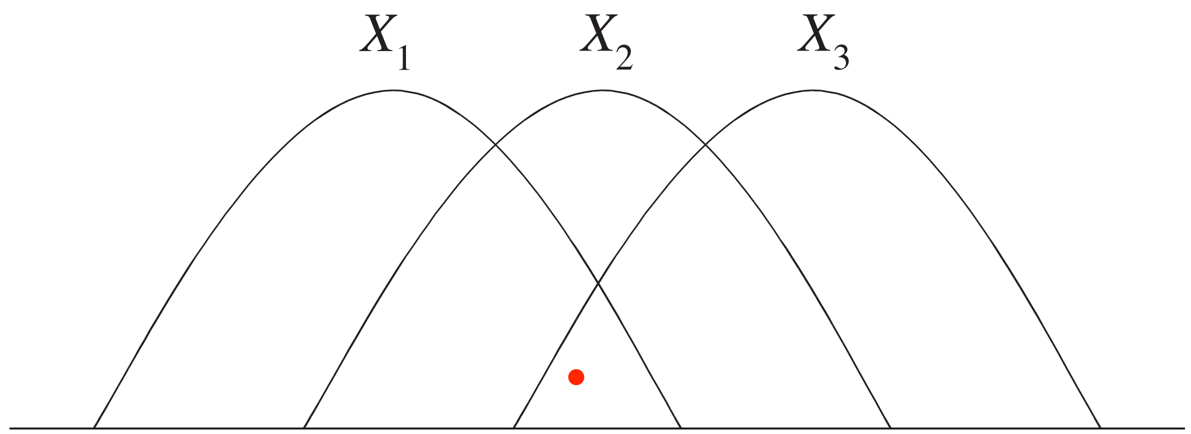
$$I(X_1; X_3|X_2) = \mu^*(\underline{\tilde{X}_1 \cap \tilde{X}_3} - \tilde{X}_2)$$



μ^* for $X_1 \to X_2 \to X_3$

1. In this information diagram,

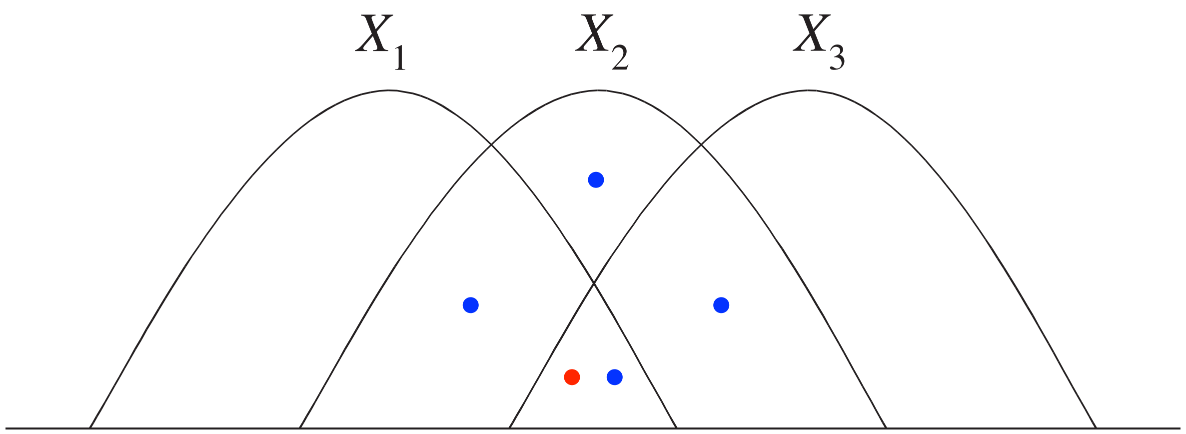
$$I(X_1; X_3 | X_2) \quad = \quad \mu^*(\tilde{X}_1 \cap \tilde{X}_3 - \tilde{X}_2)$$



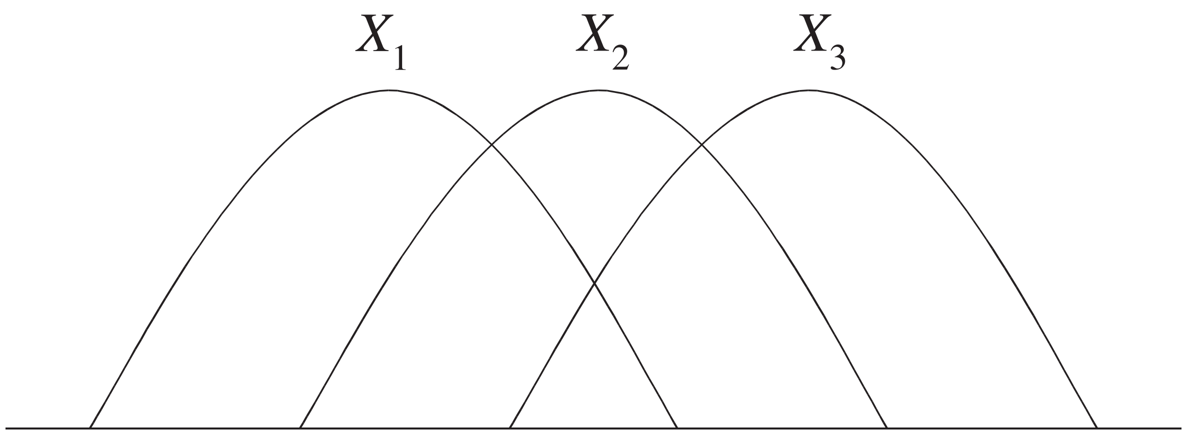
μ^* for $X_1 \to X_2 \to X_3$

1. In this information diagram,

$$I(X_1; X_3|X_2) \quad = \quad \mu^*(\underline{\tilde{X}_1 \cap \tilde{X}_3} - \underline{\tilde{X}_2})$$



μ^* for $X_1 \to X_2 \to X_3$



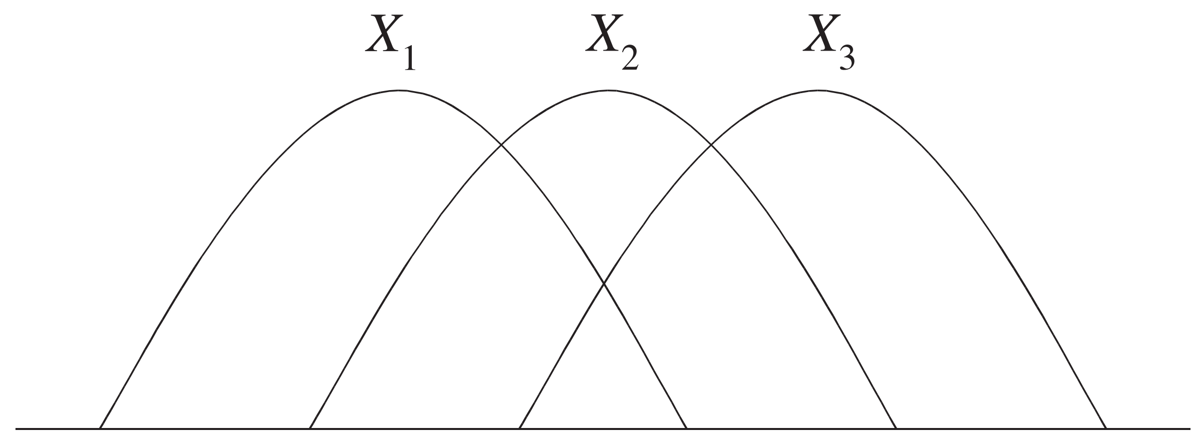
1. In this information diagram,

$$
\begin{aligned}
I(X_1; X_3 | X_2) &= \mu^*(\tilde{X}_1 \cap \tilde{X}_3 - \tilde{X}_2) \\
&= \mu^*(\emptyset)
\end{aligned}
$$

μ^* for $X_1 \to X_2 \to X_3$

1. In this information diagram,

$$
\begin{aligned}
I(X_1; X_3 | X_2) &= \mu^*(\tilde{X}_1 \cap \tilde{X}_3 - \tilde{X}_2) \\
&= \mu^*(\emptyset) \\
&= 0.
\end{aligned}
$$

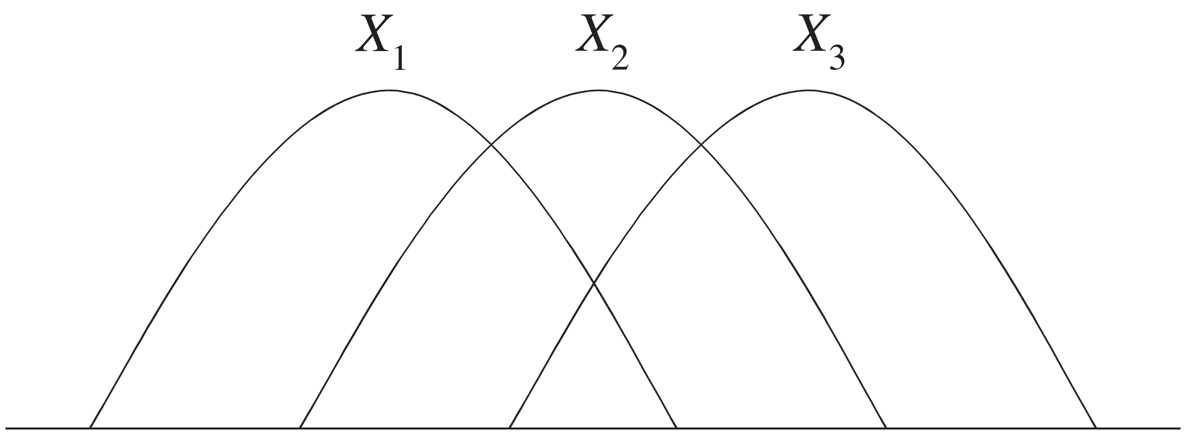


μ^* for $X_1 \to X_2 \to X_3$

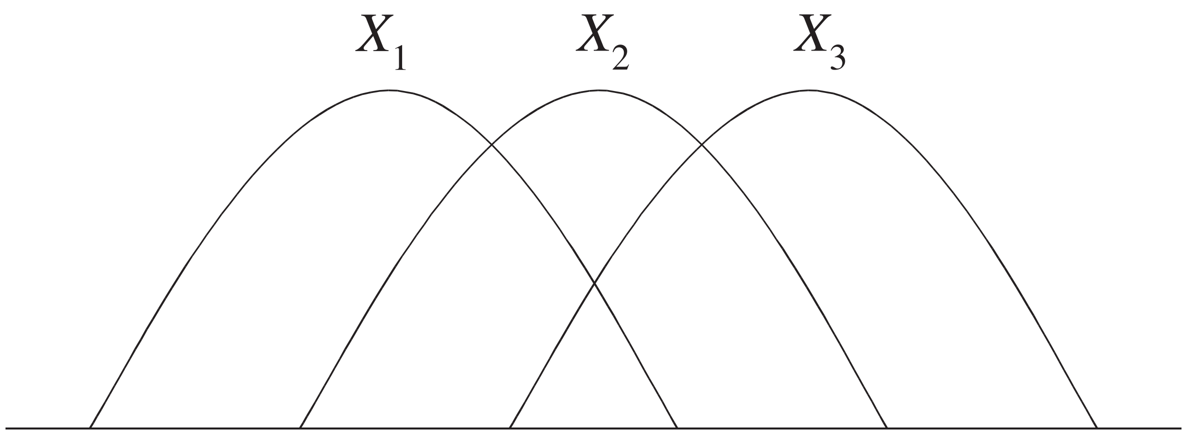
1. In this information diagram,

$$\begin{aligned} I(X_1; X_3|X_2) &= \mu^*(\tilde{X}_1 \cap \tilde{X}_3 - \tilde{X}_2) \\ &= \mu^*(\emptyset) \\ &= 0. \end{aligned}$$

2. Also,



μ^* for $X_1 \to X_2 \to X_3$



1. In this information diagram,

$$
\begin{aligned}
I(X_1; X_3|X_2) &= \mu^*(\tilde{X}_1 \cap \tilde{X}_3 - \tilde{X}_2) \\
&= \mu^*(\emptyset) \\
&= 0.
\end{aligned}
$$

2. Also,

$$\mu^*(\tilde{X}_1 \cap \tilde{X}_2 \cap \tilde{X}_3) = \mu^*(\tilde{X}_1 \cap \tilde{X}_3)$$

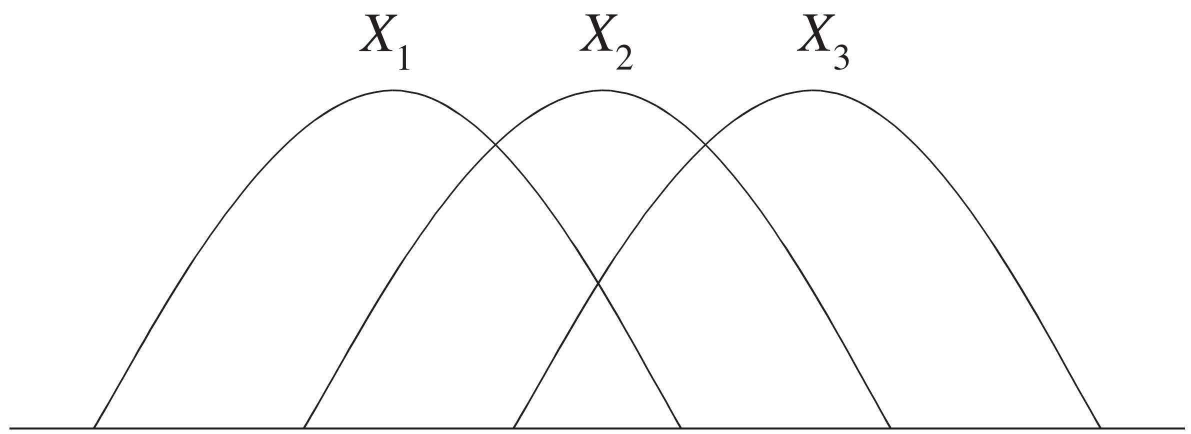
μ^* for $X_1 \rightarrow X_2 \rightarrow X_3$

1. In this information diagram,

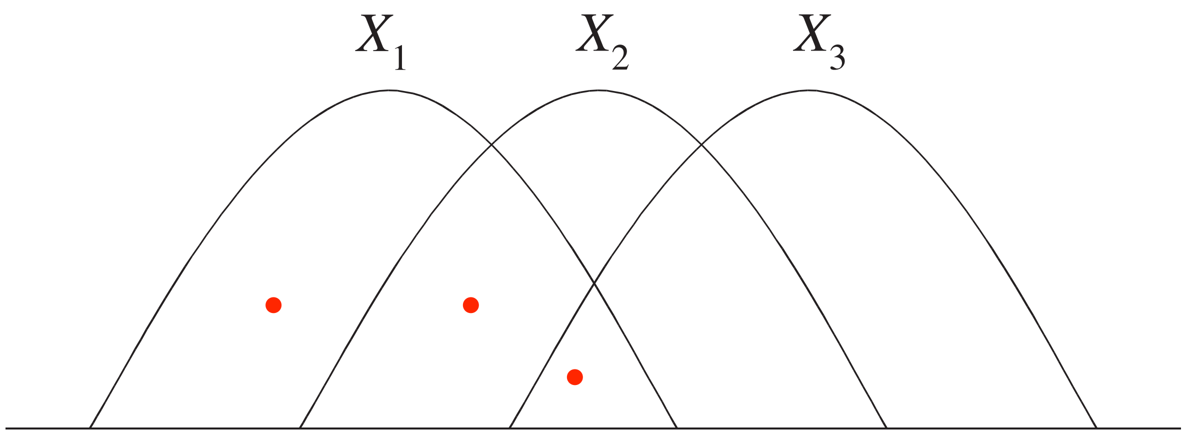
$$\begin{aligned} I(X_1; X_3|X_2) &= \mu^*(\tilde{X}_1 \cap \tilde{X}_3 - \tilde{X}_2) \\ &= \mu^*(\emptyset) \\ &= 0. \end{aligned}$$

2. Also,

$$\mu^*(\tilde{X}_1 \cap \tilde{X}_2 \cap \tilde{X}_3) = \mu^*(\tilde{X}_1 \cap \tilde{X}_3)$$



μ^* for $X_1 \to X_2 \to X_3$



1. In this information diagram,

$$\begin{aligned} I(X_1; X_3 | X_2) &= \mu^*(\tilde{X}_1 \cap \tilde{X}_3 - \tilde{X}_2) \\ &= \mu^*(\emptyset) \\ &= 0. \end{aligned}$$

2. Also,

$$\mu^*(\tilde{X}_1 \cap \tilde{X}_2 \cap \tilde{X}_3) = \mu^*(\tilde{X}_1 \cap \tilde{X}_3)$$

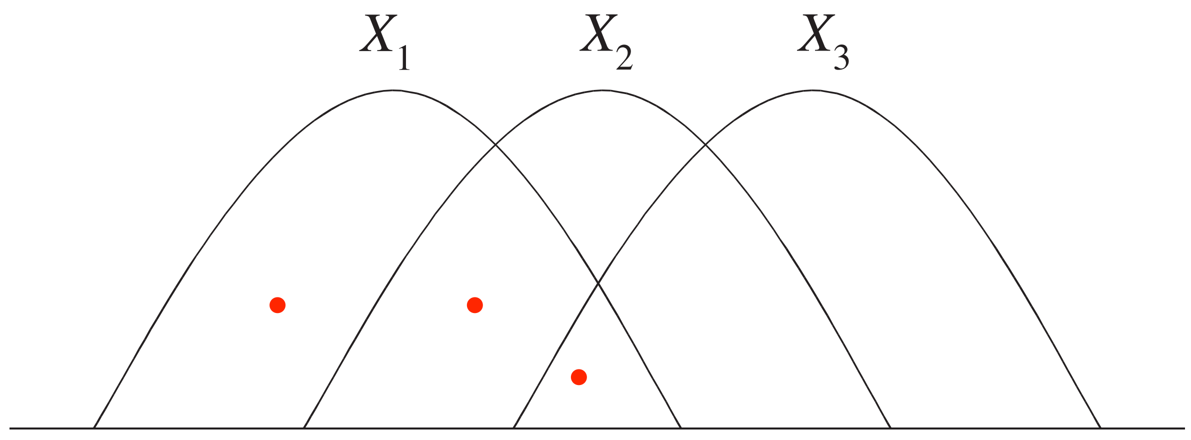
μ^* for $X_1 \to X_2 \to X_3$

1. In this information diagram,

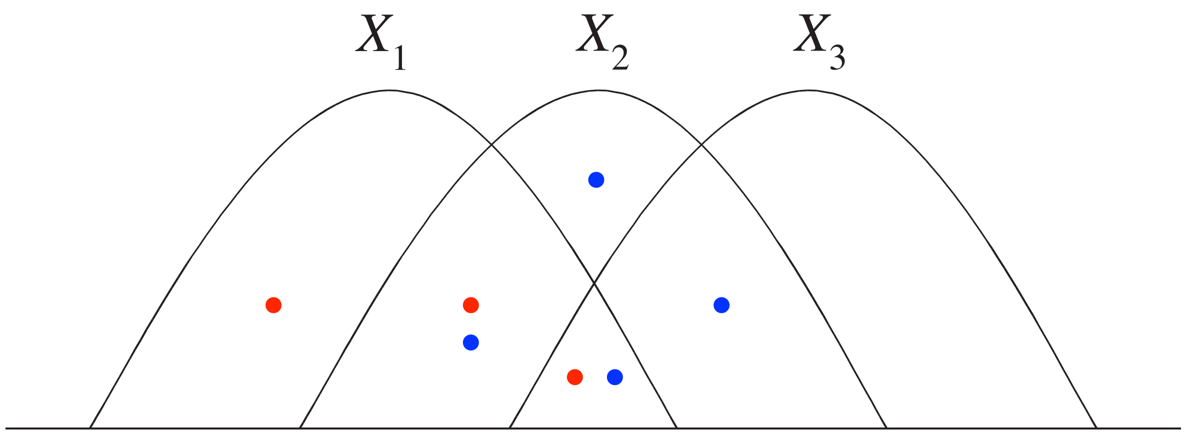
$$\begin{aligned} I(X_1; X_3|X_2) &= \mu^*(\tilde{X}_1 \cap \tilde{X}_3 - \tilde{X}_2) \\ &= \mu^*(\emptyset) \\ &= 0. \end{aligned}$$

2. Also,

$$\mu^*(\tilde{X}_1 \cap \tilde{X}_2 \cap \tilde{X}_3) = \mu^*(\tilde{X}_1 \cap \tilde{X}_3)$$



μ^* for $X_1 \to X_2 \to X_3$



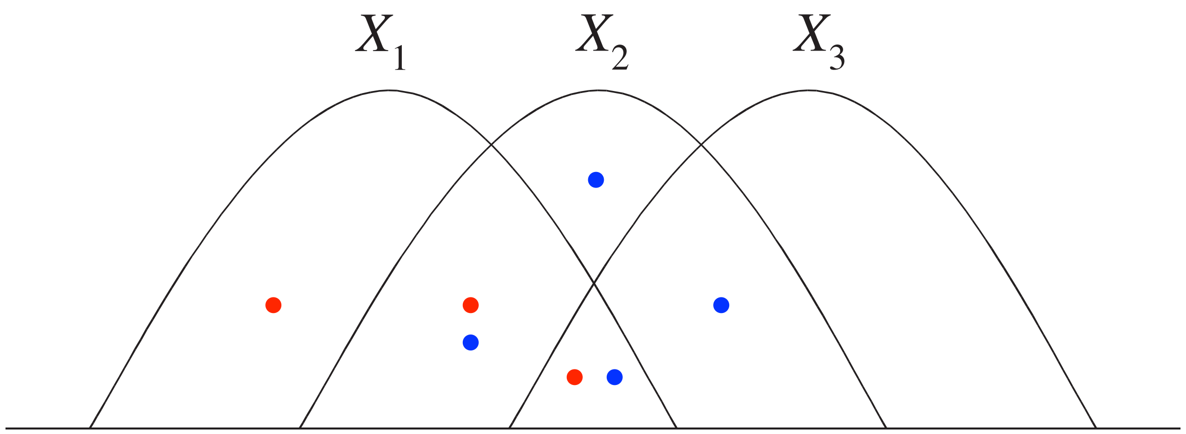
1. In this information diagram,

$$\begin{aligned} I(X_1; X_3|X_2) &= \mu^*(\tilde{X}_1 \cap \tilde{X}_3 - \tilde{X}_2) \\ &= \mu^*(\emptyset) \\ &= 0. \end{aligned}$$

2. Also,

$$\mu^*(\underline{\tilde{X}_1} \cap \underline{\tilde{X}_2} \cap \tilde{X}_3) = \mu^*(\tilde{X}_1 \cap \tilde{X}_3)$$

μ^* for $X_1 \rightarrow X_2 \rightarrow X_3$



1. In this information diagram,

$$
\begin{aligned}
I(X_1; X_3 | X_2) &= \mu^*(\tilde{X}_1 \cap \tilde{X}_3 - \tilde{X}_2) \\
&= \mu^*(\emptyset) \\
&= 0.
\end{aligned}
$$

2. Also,

$$\mu^*(\underline{\tilde{X}_1} \cap \underline{\tilde{X}_2} \cap \underline{\tilde{X}_3}) = \mu^*(\tilde{X}_1 \cap \tilde{X}_3)$$

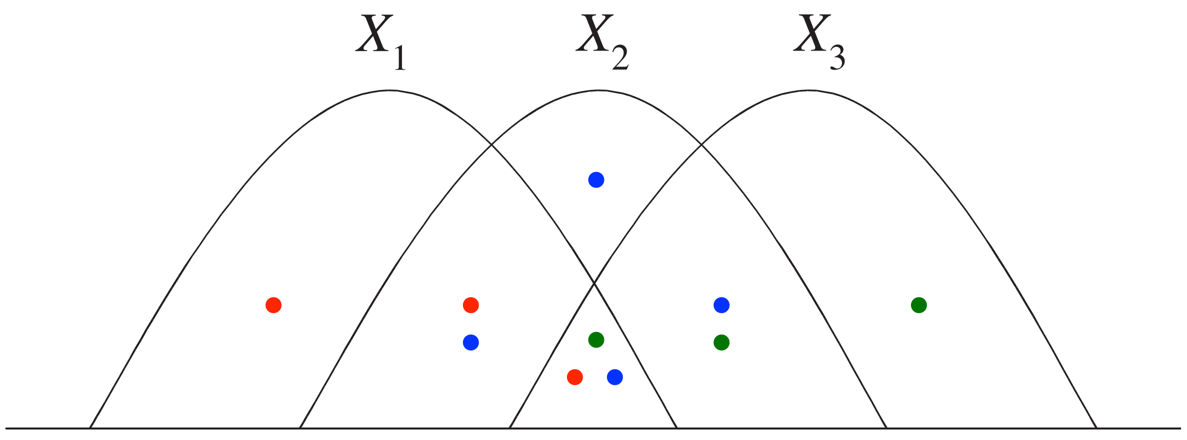
μ^* for $X_1 \rightarrow X_2 \rightarrow X_3$

1. In this information diagram,

$$\begin{aligned} I(X_1; X_3|X_2) &= \mu^*(\tilde{X}_1 \cap \tilde{X}_3 - \tilde{X}_2) \\ &= \mu^*(\emptyset) \\ &= 0. \end{aligned}$$

2. Also,

$$\mu^*(\underline{\tilde{X}_1} \cap \underline{\tilde{X}_2} \cap \underline{\tilde{X}_3}) = \mu^*(\tilde{X}_1 \cap \tilde{X}_3)$$



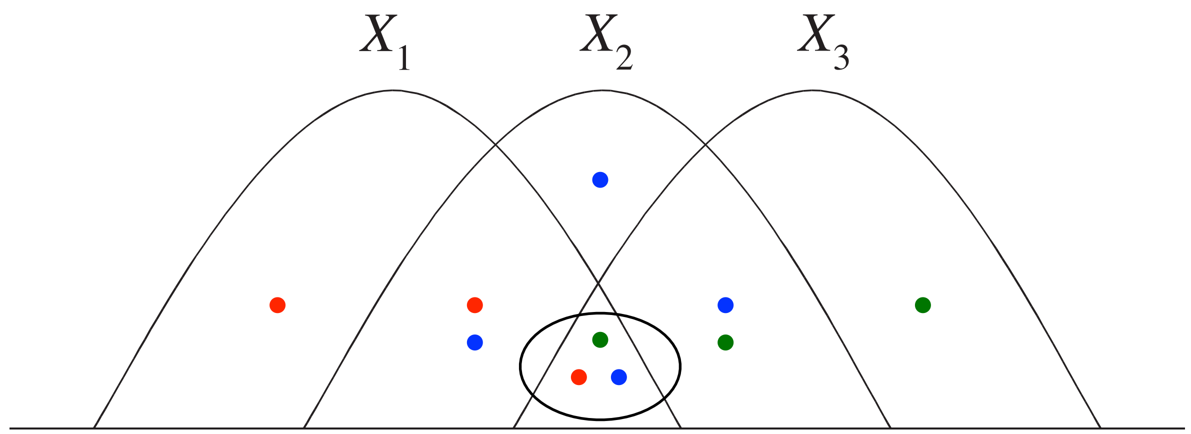
μ^* for $X_1 \to X_2 \to X_3$

1. In this information diagram,

$$\begin{aligned} I(X_1; X_3|X_2) &= \mu^*(\tilde{X}_1 \cap \tilde{X}_3 - \tilde{X}_2) \\ &= \mu^*(\emptyset) \\ &= 0. \end{aligned}$$

2. Also,

$$\mu^*(\underline{\tilde{X}_1} \cap \underline{\tilde{X}_2} \cap \underline{\tilde{X}_3}) = \mu^*(\tilde{X}_1 \cap \tilde{X}_3)$$



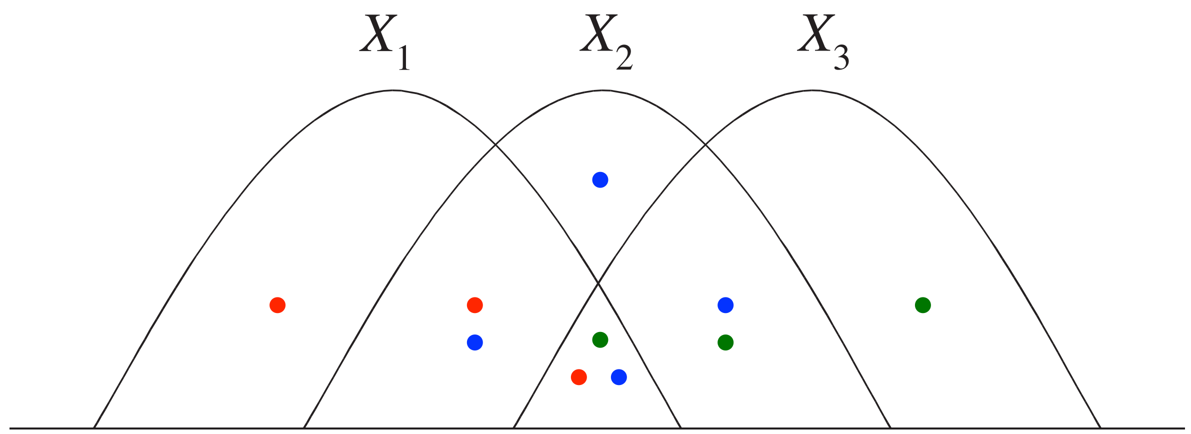
μ^* for $X_1 \to X_2 \to X_3$

1. In this information diagram,

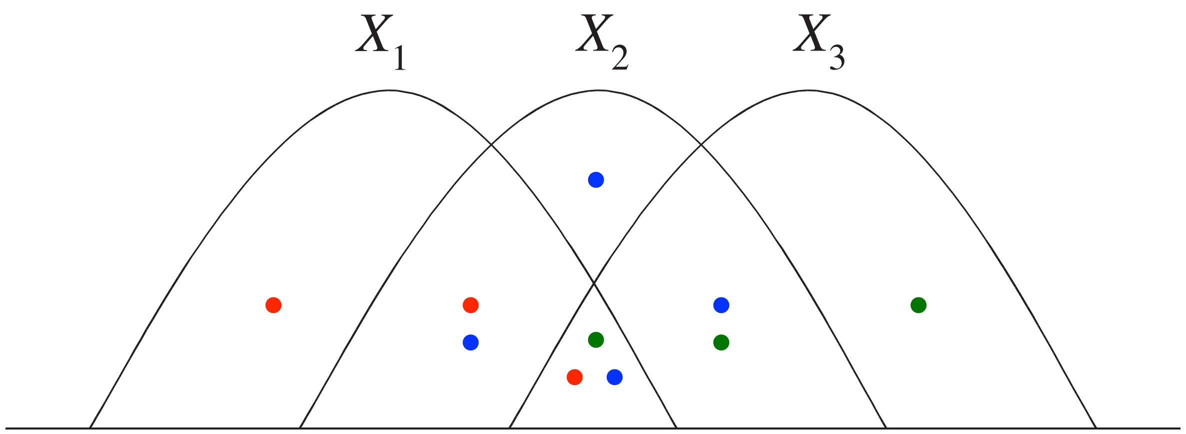
$$\begin{aligned} I(X_1; X_3 | X_2) &= \mu^*(\tilde{X}_1 \cap \tilde{X}_3 - \tilde{X}_2) \\ &= \mu^*(\emptyset) \\ &= 0. \end{aligned}$$

2. Also,

$$\mu^*(\underline{\tilde{X}_1} \cap \underline{\tilde{X}_2} \cap \underline{\tilde{X}_3}) = \mu^*(\tilde{X}_1 \cap \tilde{X}_3)$$



μ^* for $X_1 \rightarrow X_2 \rightarrow X_3$



1. In this information diagram,

$$\begin{aligned} I(X_1; X_3|X_2) &= \mu^*(\tilde{X}_1 \cap \tilde{X}_3 - \tilde{X}_2) \\ &= \mu^*(\emptyset) \\ &= 0. \end{aligned}$$

2. Also,

$$\mu^*(\underline{\tilde{X}_1} \cap \underline{\tilde{X}_2} \cap \underline{\tilde{X}_3}) = \mu^*(\underline{\tilde{X}_1} \cap \tilde{X}_3)$$

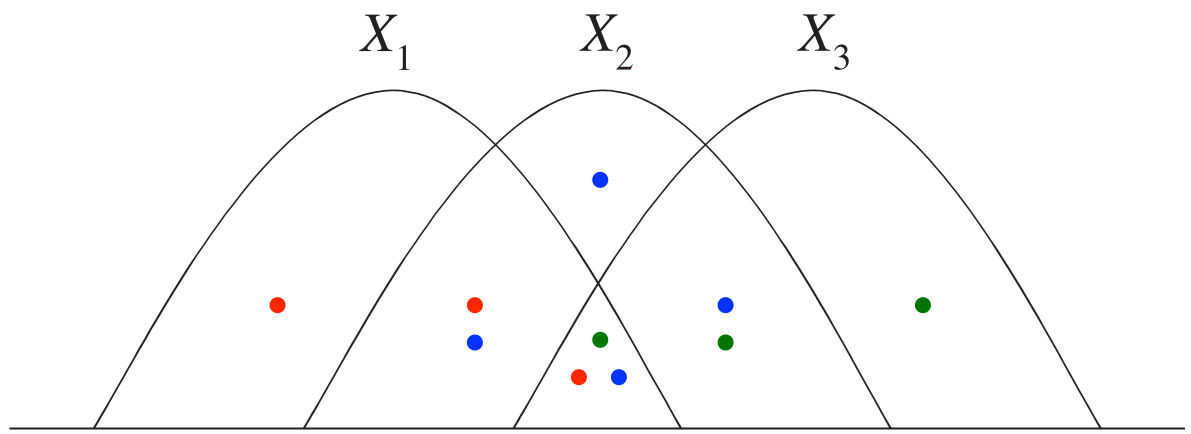
μ^* for $X_1 \to X_2 \to X_3$

1. In this information diagram,

$$\begin{aligned} I(X_1; X_3|X_2) &= \mu^*(\tilde{X}_1 \cap \tilde{X}_3 - \tilde{X}_2) \\ &= \mu^*(\emptyset) \\ &= 0. \end{aligned}$$

2. Also,

$$\mu^*(\underline{\tilde{X}_1} \cap \underline{\tilde{X}_2} \cap \underline{\tilde{X}_3}) = \mu^*(\underline{\tilde{X}_1} \cap \underline{\tilde{X}_3})$$



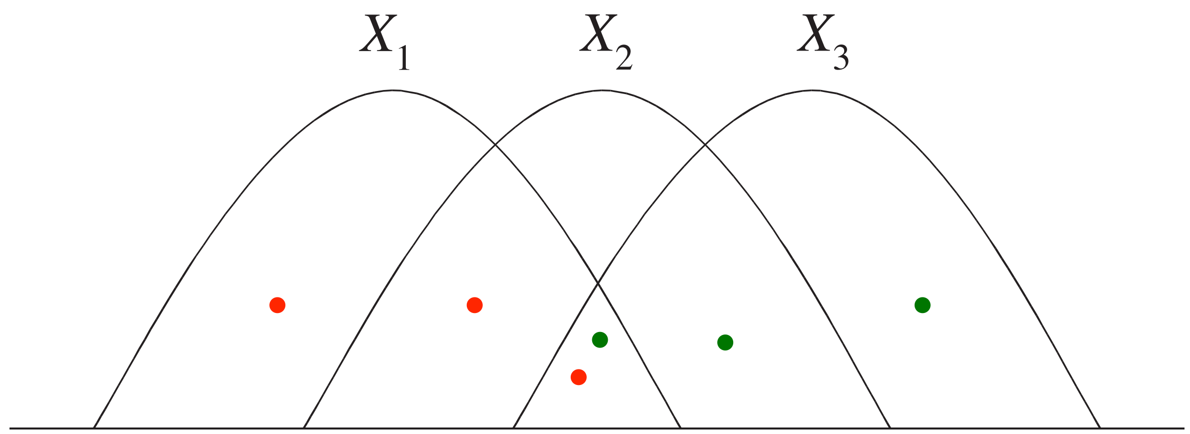
μ^* for $X_1 \to X_2 \to X_3$

1. In this information diagram,

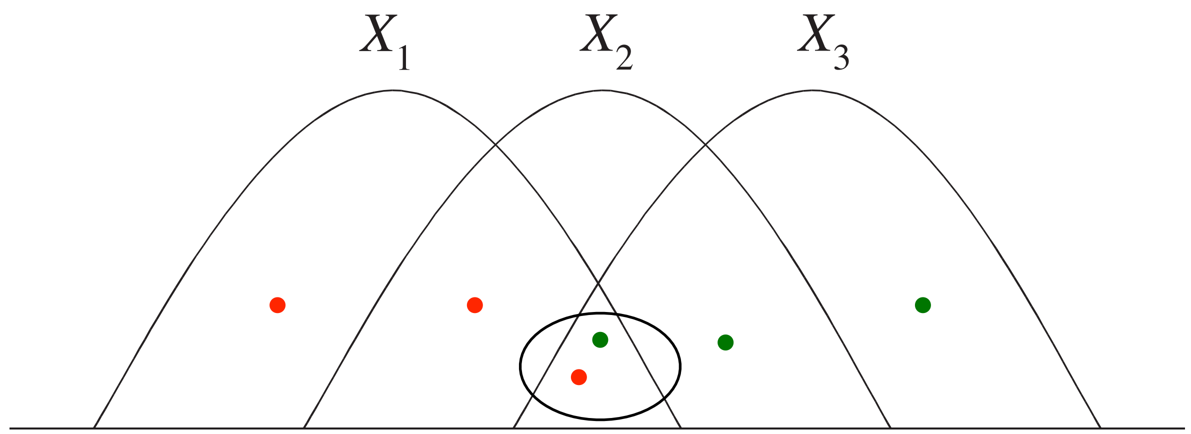
$$I(X_1; X_3 | X_2) = \mu^*(\tilde{X}_1 \cap \tilde{X}_3 - \tilde{X}_2) = \mu^*(\emptyset) = 0.$$

2. Also,

$$\mu^*(\underline{\tilde{X}_1} \cap \tilde{X}_2 \cap \underline{\tilde{X}_3}) = \mu^*(\underline{\tilde{X}_1} \cap \underline{\tilde{X}_3})$$



μ^* for $X_1 \rightarrow X_2 \rightarrow X_3$



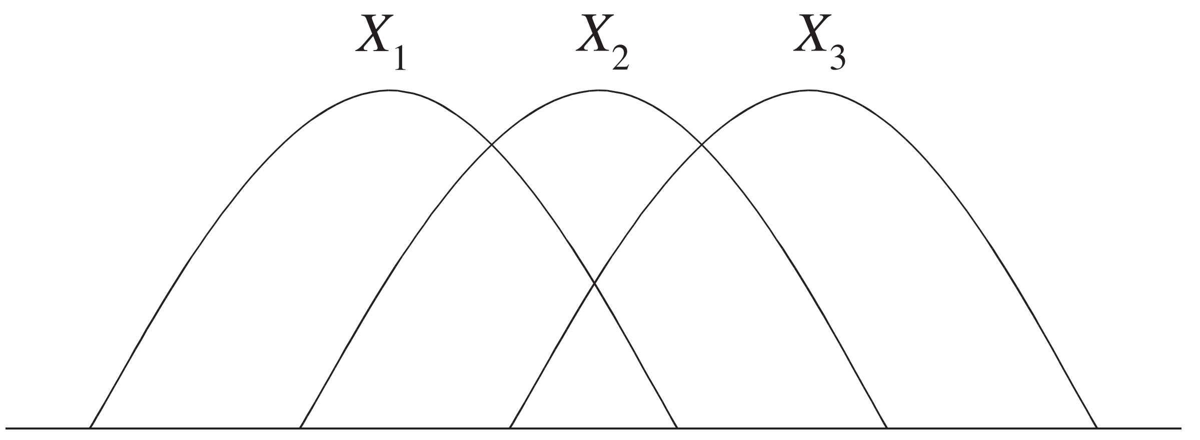
1. In this information diagram,

$$\begin{aligned} I(X_1; X_3 | X_2) &= \mu^*(\tilde{X}_1 \cap \tilde{X}_3 - \tilde{X}_2) \\ &= \mu^*(\emptyset) \\ &= 0. \end{aligned}$$

2. Also,

$$\mu^*(\underline{\tilde{X}_1} \cap \tilde{X}_2 \cap \underline{\tilde{X}_3}) = \mu^*(\underline{\tilde{X}_1} \cap \underline{\tilde{X}_3})$$

μ^* for $X_1 \to X_2 \to X_3$



1. In this information diagram,

$$
\begin{aligned}
I(X_1; X_3|X_2) &= \mu^*(\tilde{X}_1 \cap \tilde{X}_3 - \tilde{X}_2) \\
&= \mu^*(\emptyset) \\
&= 0.
\end{aligned}
$$

2. Also,

$$
\begin{aligned}
\mu^*(\tilde{X}_1 \cap \tilde{X}_2 \cap \tilde{X}_3) &= \mu^*(\tilde{X}_1 \cap \tilde{X}_3) \\
&= I(X_1; X_3)
\end{aligned}
$$

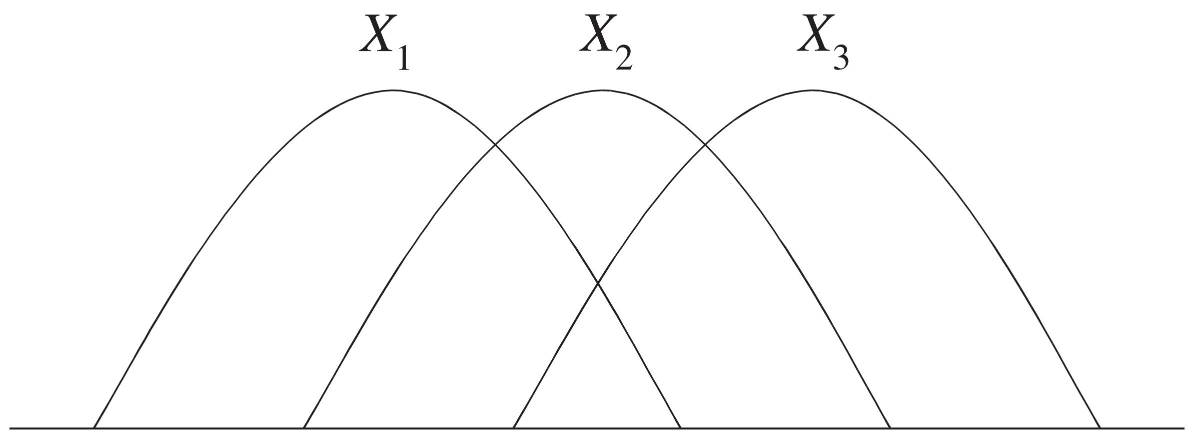
μ^* for $X_1 \to X_2 \to X_3$

1. In this information diagram,

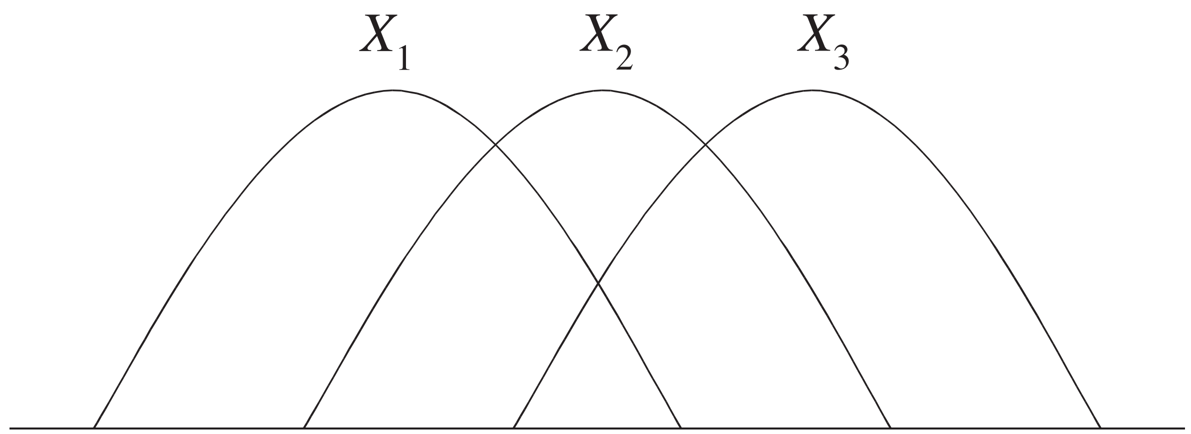
$$\begin{aligned} I(X_1;X_3|X_2) &= \mu^*(\tilde{X}_1 \cap \tilde{X}_3 - \tilde{X}_2) \\ &= \mu^*(\emptyset) \\ &= 0. \end{aligned}$$

2. Also,

$$\begin{aligned} \mu^*(\tilde{X}_1 \cap \tilde{X}_2 \cap \tilde{X}_3) &= \mu^*(\tilde{X}_1 \cap \tilde{X}_3) \\ &= I(X_1;X_3) \\ &\geq 0. \end{aligned}$$



μ^* for $X_1 \to X_2 \to X_3$



1. In this information diagram,

$$\begin{aligned} I(X_1; X_3|X_2) &= \mu^*(\tilde{X}_1 \cap \tilde{X}_3 - \tilde{X}_2) \\ &= \mu^*(\emptyset) \\ &= 0. \end{aligned}$$

2. Also,

$$\begin{aligned} \mu^*(\tilde{X}_1 \cap \tilde{X}_2 \cap \tilde{X}_3) &= \mu^*(\tilde{X}_1 \cap \tilde{X}_3) \\ &= I(X_1; X_3) \\ &\geq 0. \end{aligned}$$

3. Since the values of μ^* on all the remaining atoms correspond to Shannon's information measures and hence are nonnegative, we conclude that μ^* is a measure.

- For $X_1 \to X_2 \to X_3 \to X_4$, μ^* vanishes on the following 5 atoms:

$$\tilde{X}_1 \cap \tilde{X}_2^c \cap \tilde{X}_3 \cap \tilde{X}_4^c$$
$$\tilde{X}_1 \cap \tilde{X}_2^c \cap \tilde{X}_3 \cap \tilde{X}_4$$
$$\tilde{X}_1 \cap \tilde{X}_2^c \cap \tilde{X}_3^c \cap \tilde{X}_4$$
$$\tilde{X}_1 \cap \tilde{X}_2 \cap \tilde{X}_3^c \cap \tilde{X}_4$$
$$\tilde{X}_1^c \cap \tilde{X}_2 \cap \tilde{X}_3^c \cap \tilde{X}_4$$

- For $X_1 \to X_2 \to X_3 \to X_4$, μ^* vanishes on the following 5 atoms:

$$\tilde{X}_1 \cap \tilde{X}_2^c \cap \tilde{X}_3 \cap \tilde{X}_4^c$$
$$\tilde{X}_1 \cap \tilde{X}_2^c \cap \tilde{X}_3 \cap \tilde{X}_4$$
$$\tilde{X}_1 \cap \tilde{X}_2^c \cap \tilde{X}_3^c \cap \tilde{X}_4$$
$$\tilde{X}_1 \cap \tilde{X}_2 \cap \tilde{X}_3^c \cap \tilde{X}_4$$
$$\tilde{X}_1^c \cap \tilde{X}_2 \cap \tilde{X}_3^c \cap \tilde{X}_4$$

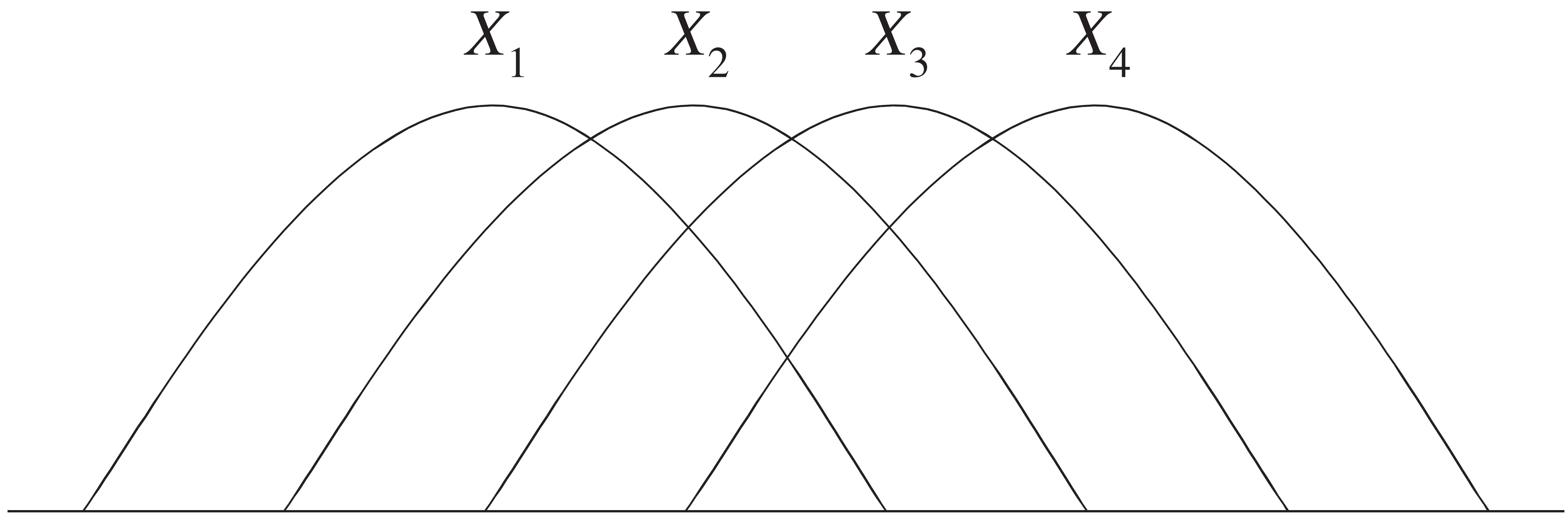
- The information diagram can be displayed in two dimensions.

- For $X_1 \to X_2 \to X_3 \to X_4$, μ^* vanishes on the following 5 atoms:

$$\tilde{X}_1 \cap \tilde{X}_2^c \cap \tilde{X}_3 \cap \tilde{X}_4^c$$
$$\tilde{X}_1 \cap \tilde{X}_2^c \cap \tilde{X}_3 \cap \tilde{X}_4$$
$$\tilde{X}_1 \cap \tilde{X}_2^c \cap \tilde{X}_3^c \cap \tilde{X}_4$$
$$\tilde{X}_1 \cap \tilde{X}_2 \cap \tilde{X}_3^c \cap \tilde{X}_4$$
$$\tilde{X}_1^c \cap \tilde{X}_2 \cap \tilde{X}_3^c \cap \tilde{X}_4$$

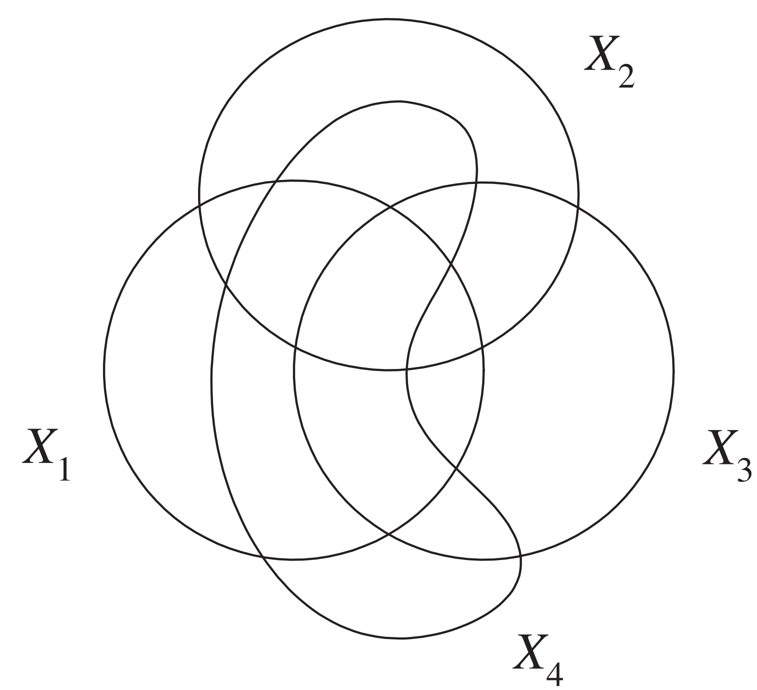
- The information diagram can be displayed in two dimensions.

- The values of μ^* on the remaining atoms correspond to Shannon's information measures and hence are nonnegative. Thus, μ^* is a measure.



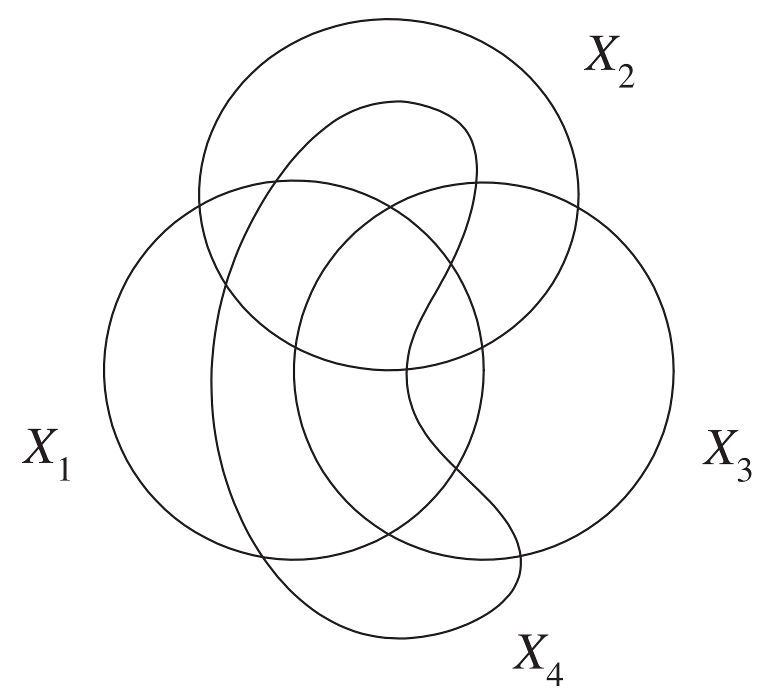
X1
X2
X3
X4

Structure of μ^* for $X_1 \to X_2 \to X_3 \to X_4$



Structure of μ^* for $X_1 \to X_2 \to X_3 \to X_4$

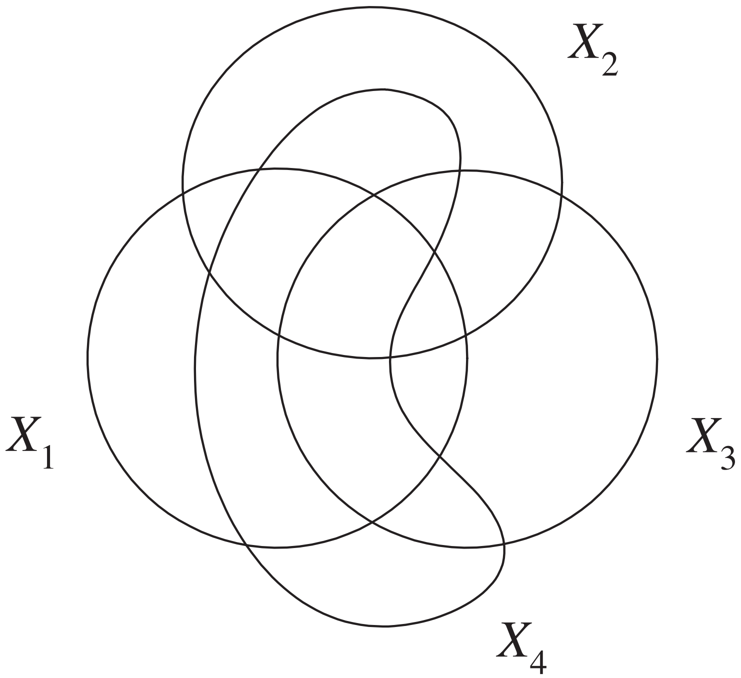
1. The Markov subchain $X_1 \to X_2 \to X_3$ implies



Structure of μ^* for $X_1 \to X_2 \to X_3 \to X_4$

1. The Markov subchain $X_1 \to X_2 \to X_3$ implies

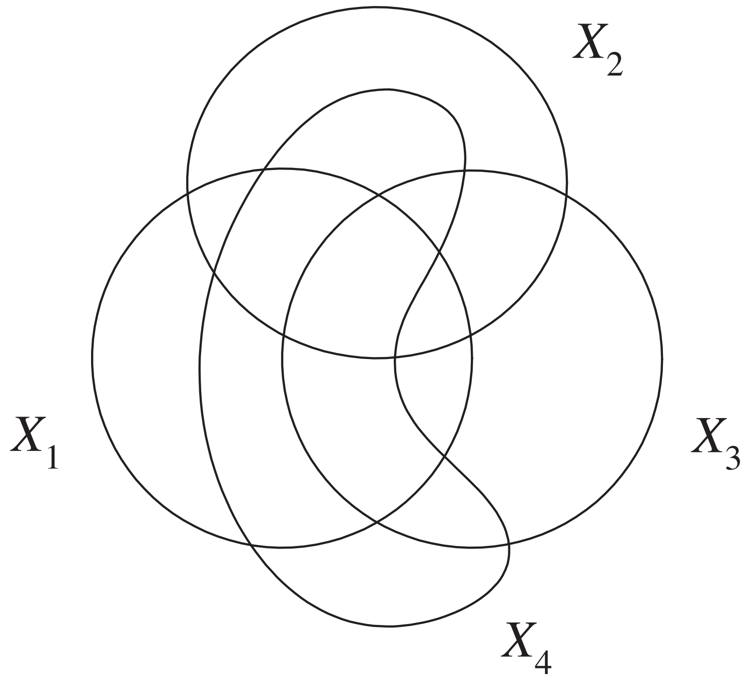
$$0 = I(X_1; X_3|X_2) = I(X_1; X_3; X_4|X_2) + I(X_1; X_3|X_2, X_4).$$



Structure of μ^* for $X_1 \rightarrow X_2 \rightarrow X_3 \rightarrow X_4$

1. The Markov subchain $X_1 \rightarrow X_2 \rightarrow X_3$ implies

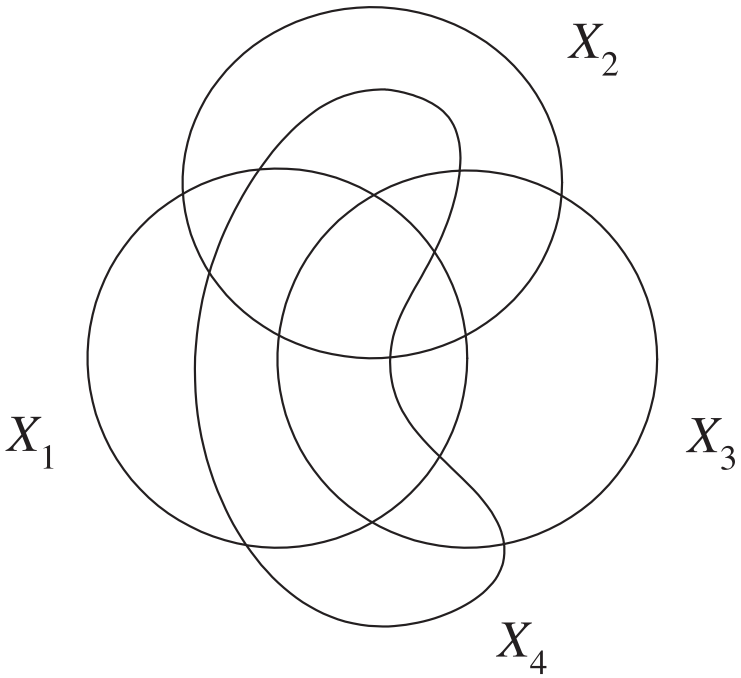
$$0 = I(\underline{X_1; X_3}|X_2) = I(\underline{X_1; X_3}; X_4|X_2) + I(\underline{X_1; X_3}|X_2, X_4).$$



Structure of μ^* for $X_1 \to X_2 \to X_3 \to X_4$

1. The Markov subchain $X_1 \to X_2 \to X_3$ implies

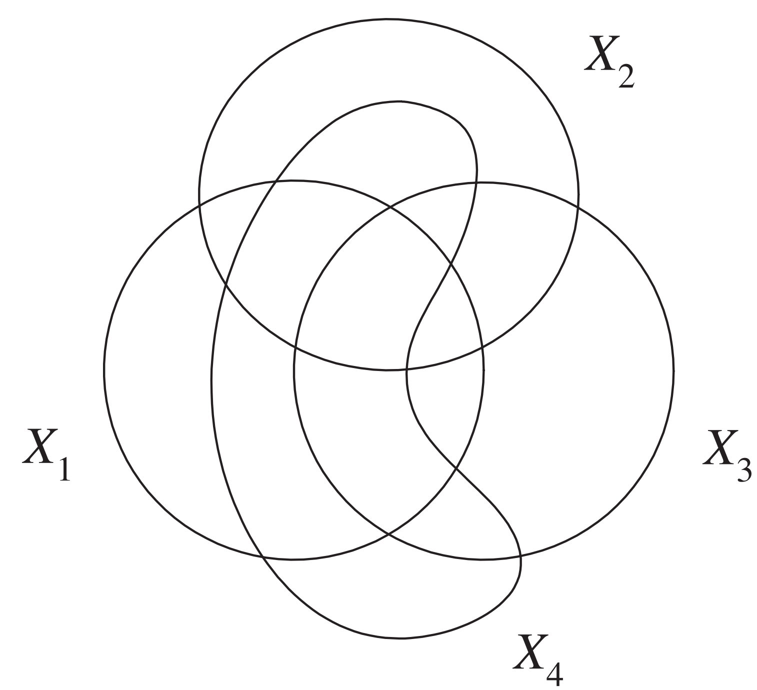
$$0 = I(X_1; X_3|X_2) = I(X_1; X_3; X_4|X_2) + I(X_1; X_3|X_2, X_4).$$



Structure of μ^* for $X_1 \to X_2 \to X_3 \to X_4$

1. The Markov subchain $X_1 \to X_2 \to X_3$ implies

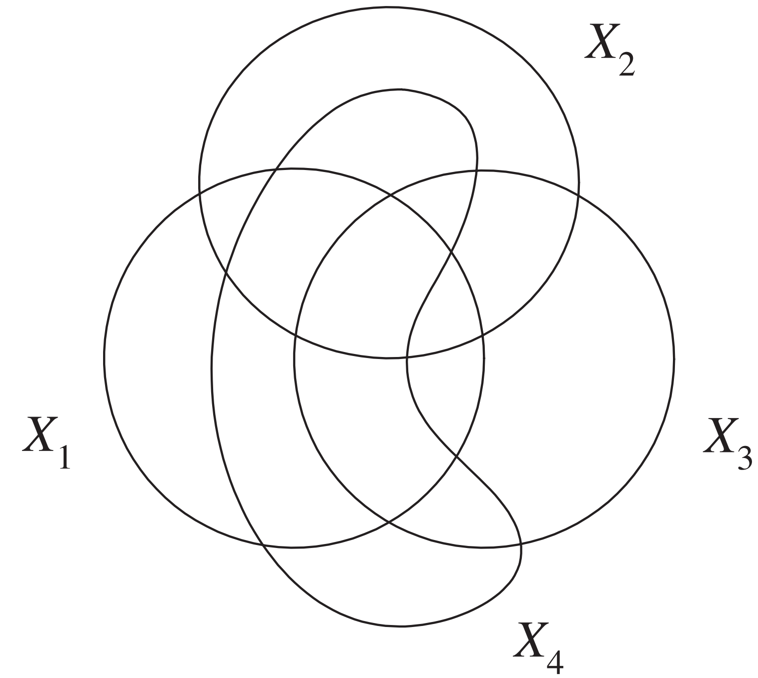
$$0 = I(\underline{X_1; X_3}|\underline{X_2}) = I(\underline{X_1; X_3}; \underline{X_4}|\underline{X_2}) + I(\underline{X_1; X_3}|\underline{X_2}, X_4).$$



Structure of μ^* for $X_1 \to X_2 \to X_3 \to X_4$

1. The Markov subchain $X_1 \to X_2 \to X_3$ implies

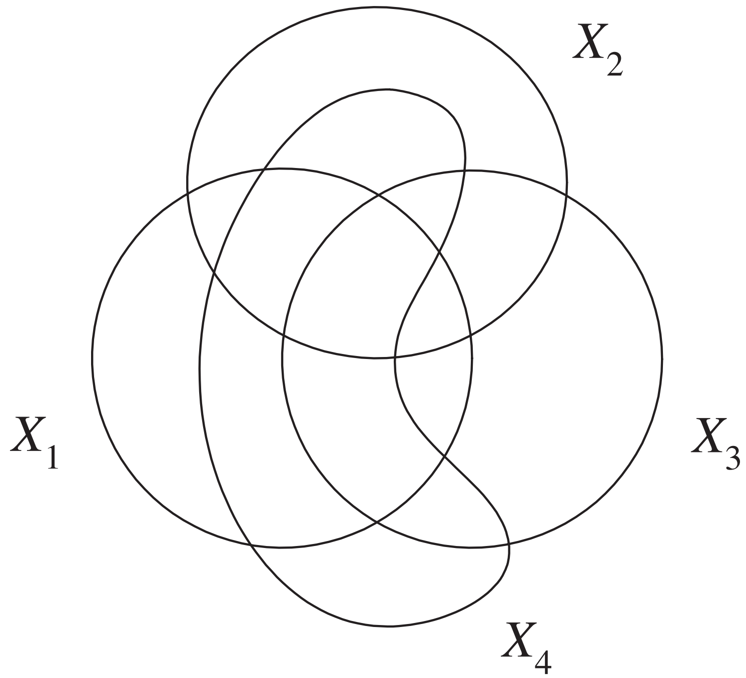
$$0 = I(X_1; X_3|X_2) = I(X_1; X_3; X_4|X_2) + I(X_1; X_3|X_2, X_4).$$



Structure of μ^* for $X_1 \to X_2 \to X_3 \to X_4$

1. The Markov subchain $X_1 \to X_2 \to X_3$ implies

$$0 = I(X_1; X_3|X_2) = I(X_1; X_3; X_4|X_2) + I(X_1; X_3|X_2, X_4).$$

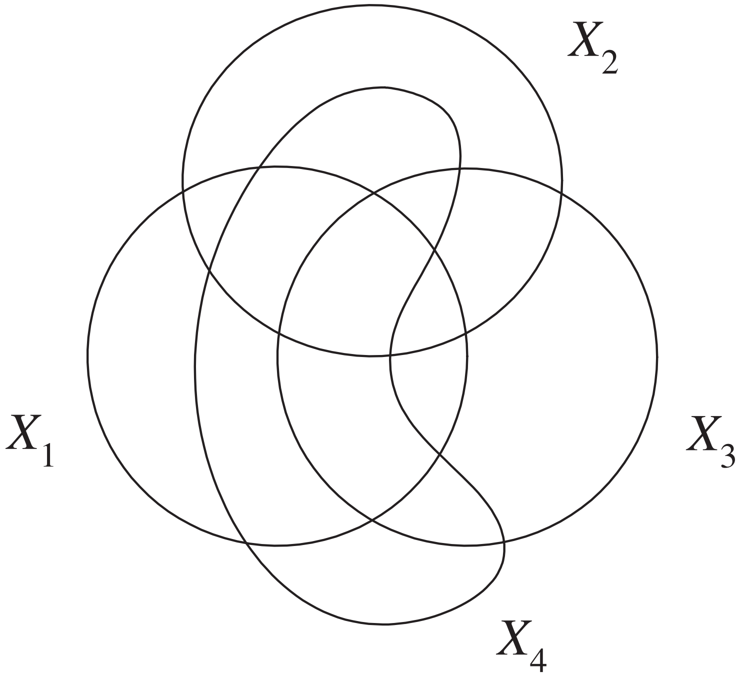


Structure of μ^* for $X_1 \to X_2 \to X_3 \to X_4$

1. The Markov subchain $X_1 \to X_2 \to X_3$ implies

$0 = I(X_1; X_3|X_2) = I(X_1; X_3; X_4|X_2) + I(X_1; X_3|X_2, X_4)$.

Let $I(X_1; X_3|X_2, X_4) = a \geq 0$. Then



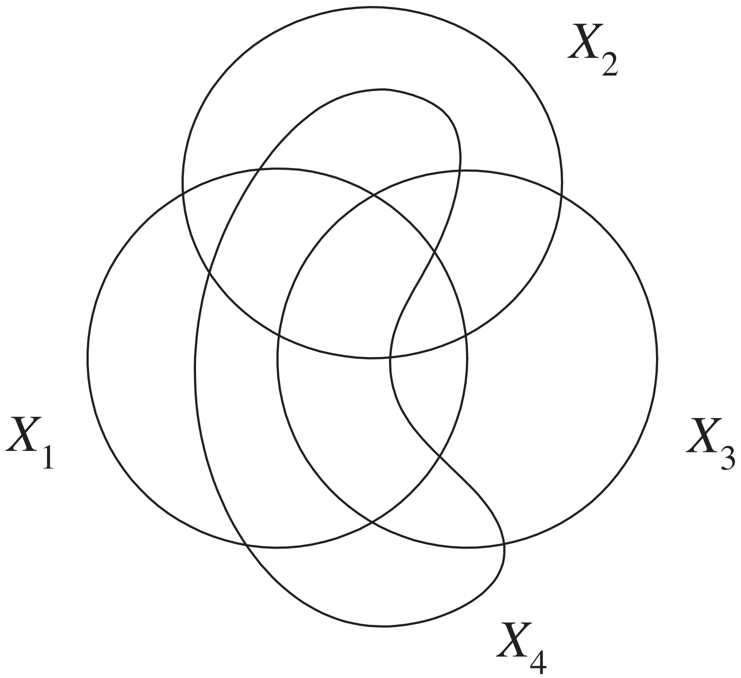
Structure of μ^* for $X_1 \to X_2 \to X_3 \to X_4$

1. The Markov subchain $X_1 \to X_2 \to X_3$ implies

$$0 = I(X_1; X_3|X_2) = \underline{I(X_1; X_3; X_4|X_2)} + \underline{I(X_1; X_3|X_2, X_4)}.$$

Let $\underline{I(X_1; X_3|X_2, X_4) = a} \geq 0$. Then

$$\underline{I(X_1; X_3; X_4|X_2) = -a}.$$



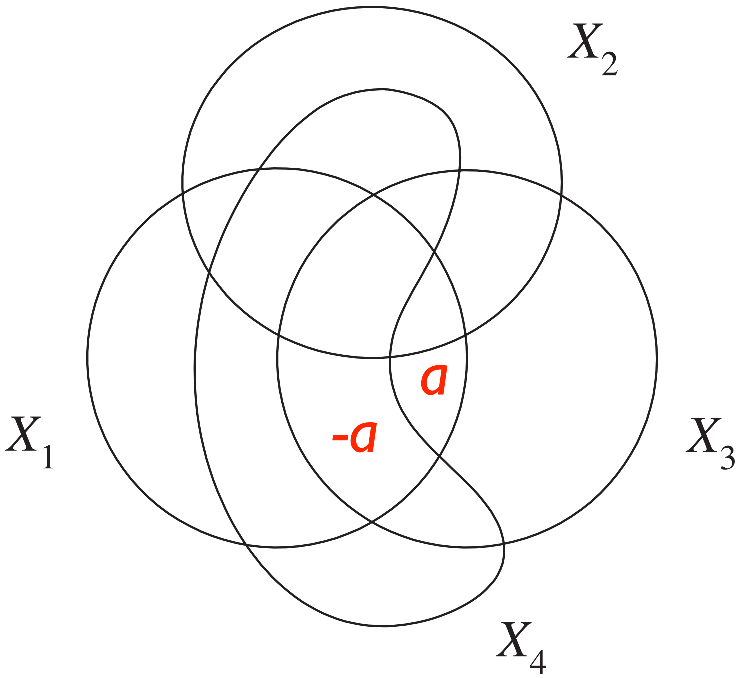
Structure of μ^* for $X_1 \to X_2 \to X_3 \to X_4$

1. The Markov subchain $X_1 \to X_2 \to X_3$ implies

$$0 = I(X_1; X_3|X_2) = \underline{I(X_1; X_3; X_4|X_2)} + \underline{I(X_1; X_3|X_2, X_4)}.$$

Let $\underline{I(X_1; X_3|X_2, X_4) = a} \geq 0$. Then

$$\underline{I(X_1; X_3; X_4|X_2) = -a}.$$



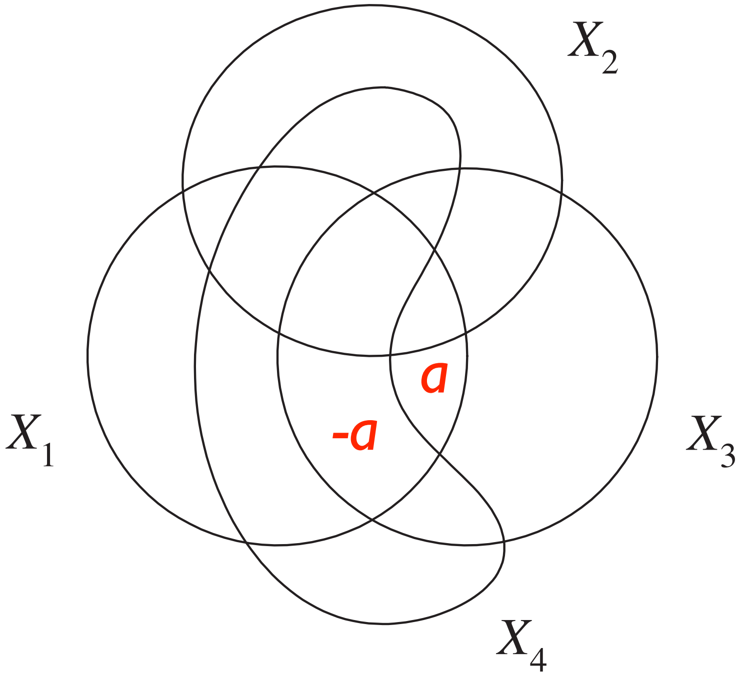
Structure of μ^* for $X_1 \to X_2 \to X_3 \to X_4$

1. The Markov subchain $X_1 \to X_2 \to X_3$ implies

$0 = I(X_1; X_3|X_2) = I(X_1; X_3; X_4|X_2) + I(X_1; X_3|X_2, X_4).$

Let $I(X_1; X_3|X_2, X_4) = a \geq 0$. Then

$$I(X_1; X_3; X_4|X_2) = -a.$$



Structure of μ^* for $X_1 \to X_2 \to X_3 \to X_4$

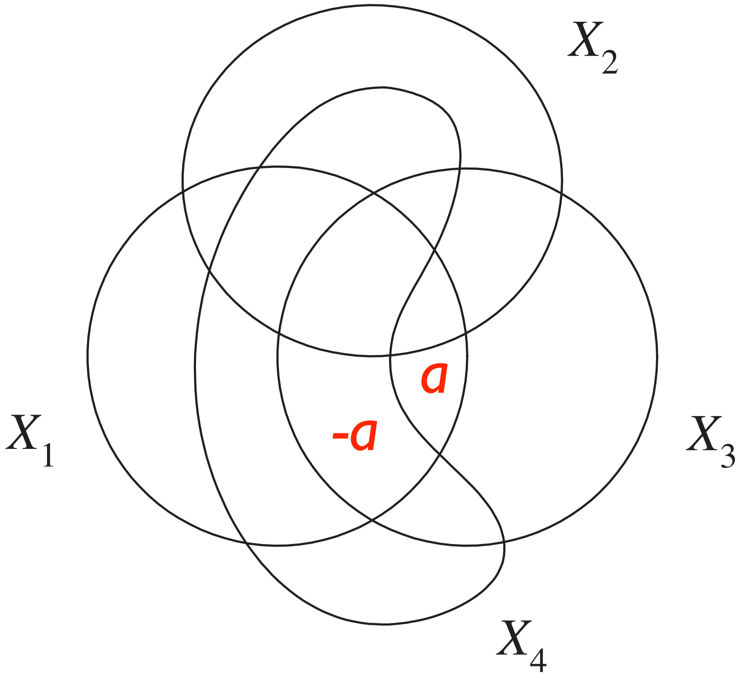
1. The Markov subchain $X_1 \to X_2 \to X_3$ implies

$$0 = I(X_1; X_3|X_2) = I(X_1; X_3; X_4|X_2) + I(X_1; X_3|X_2, X_4).$$

Let $I(X_1; X_3|X_2, X_4) = a \geq 0$. Then

$$I(X_1; X_3; X_4|X_2) = -a.$$

2. The Markov subchain $X_1 \to X_2 \to X_4$ implies



Structure of μ^* for $X_1 \to X_2 \to X_3 \to X_4$

1. The Markov subchain $X_1 \to X_2 \to X_3$ implies

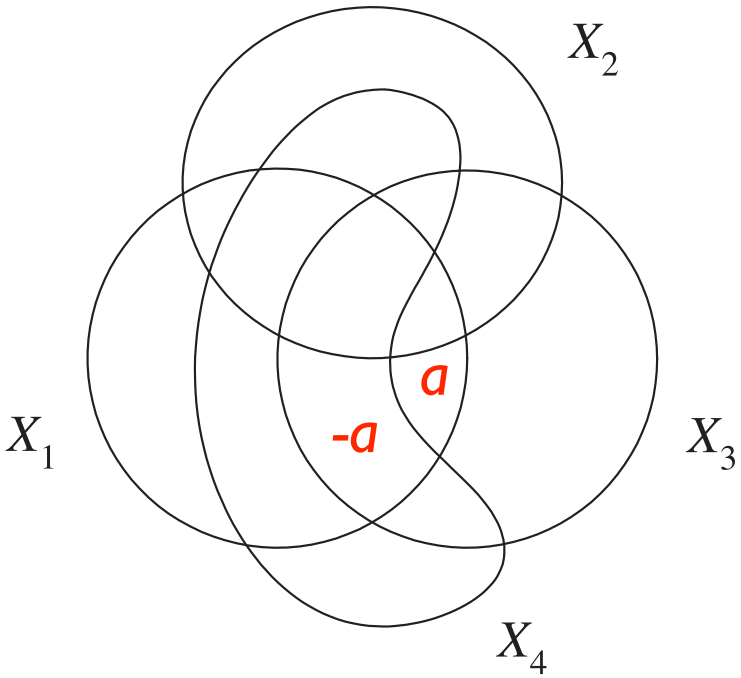
$0 = I(X_1; X_3|X_2) = I(X_1; X_3; X_4|X_2) + I(X_1; X_3|X_2, X_4)$.

Let $I(X_1; X_3|X_2, X_4) = a \geq 0$. Then

$$I(X_1; X_3; X_4|X_2) = -a.$$

2. The Markov subchain $X_1 \to X_2 \to X_4$ implies

$0 = I(X_1; X_4|X_2) = I(X_1; X_3; X_4|X_2) + I(X_1; X_4|X_2, X_3)$.



Structure of μ^* for $X_1 \to X_2 \to X_3 \to X_4$

1. The Markov subchain $X_1 \to X_2 \to X_3$ implies

$0 = I(X_1; X_3|X_2) = I(X_1; X_3; X_4|X_2) + I(X_1; X_3|X_2, X_4)$.

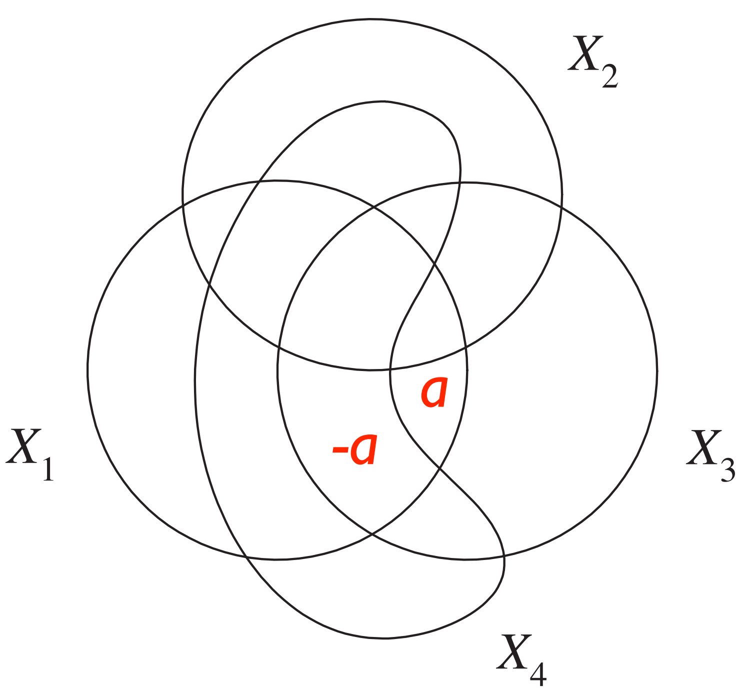
Let $I(X_1; X_3|X_2, X_4) = a \geq 0$. Then

$$I(X_1; X_3; X_4|X_2) = -a.$$

2. The Markov subchain $X_1 \to X_2 \to X_4$ implies

$0 = I(X_1; X_4|X_2) = I(X_1; X_3; X_4|X_2) + I(X_1; X_4|X_2, X_3)$.

Since $I(X_1; X_3; X_4|X_2) = -a$,



Structure of μ^* for $X_1 \to X_2 \to X_3 \to X_4$

1. The Markov subchain $X_1 \to X_2 \to X_3$ implies

$$0 = I(X_1; X_3|X_2) = I(X_1; X_3; X_4|X_2) + I(X_1; X_3|X_2, X_4).$$

Let $I(X_1; X_3|X_2, X_4) = a \geq 0$. Then

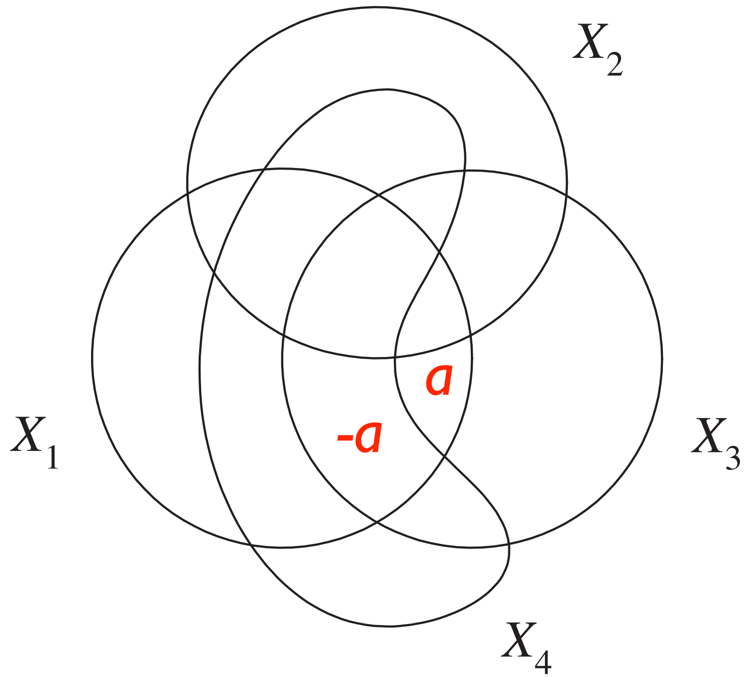
$$I(X_1; X_3; X_4|X_2) = -a.$$

2. The Markov subchain $X_1 \to X_2 \to X_4$ implies

$$0 = I(X_1; X_4|X_2) = I(X_1; X_3; X_4|X_2) + I(X_1; X_4|X_2, X_3).$$

Since $I(X_1; X_3; X_4|X_2) = -a$,

$$I(X_1; X_4|X_2, X_3) = a.$$



Structure of μ^* for $X_1 \to X_2 \to X_3 \to X_4$

1. The Markov subchain $X_1 \to X_2 \to X_3$ implies

$$0 = I(X_1; X_3|X_2) = I(X_1; X_3; X_4|X_2) + I(X_1; X_3|X_2, X_4).$$

Let $I(X_1; X_3|X_2, X_4) = a \geq 0$. Then

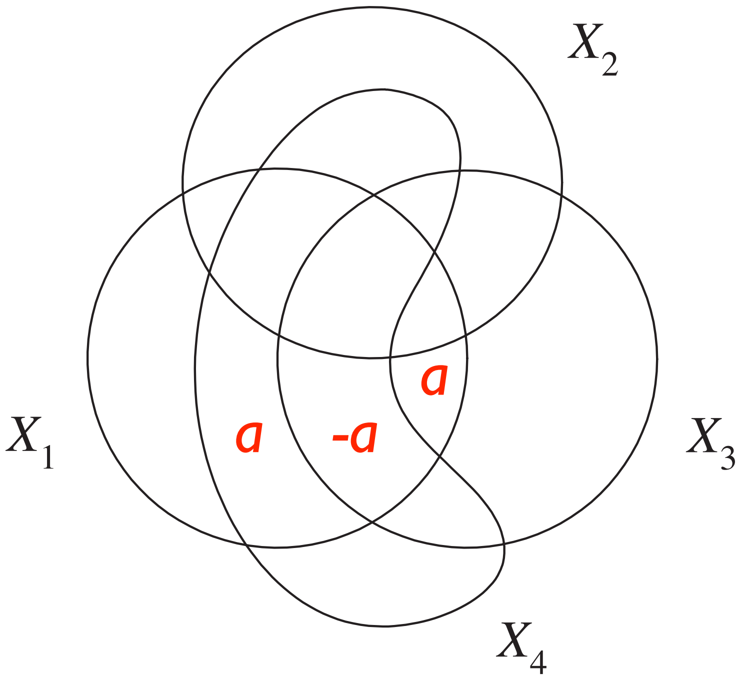
$$I(X_1; X_3; X_4|X_2) = -a.$$

2. The Markov subchain $X_1 \to X_2 \to X_4$ implies

$$0 = I(X_1; X_4|X_2) = I(X_1; X_3; X_4|X_2) + I(X_1; X_4|X_2, X_3).$$

Since $I(X_1; X_3; X_4|X_2) = -a$,

$$I(X_1; X_4|X_2, X_3) = a.$$



Structure of μ^* for $X_1 \to X_2 \to X_3 \to X_4$

1. The Markov subchain $X_1 \to X_2 \to X_3$ implies

$$0 = I(X_1; X_3|X_2) = I(X_1; X_3; X_4|X_2) + I(X_1; X_3|X_2, X_4).$$

Let $I(X_1; X_3|X_2, X_4) = a \geq 0$. Then

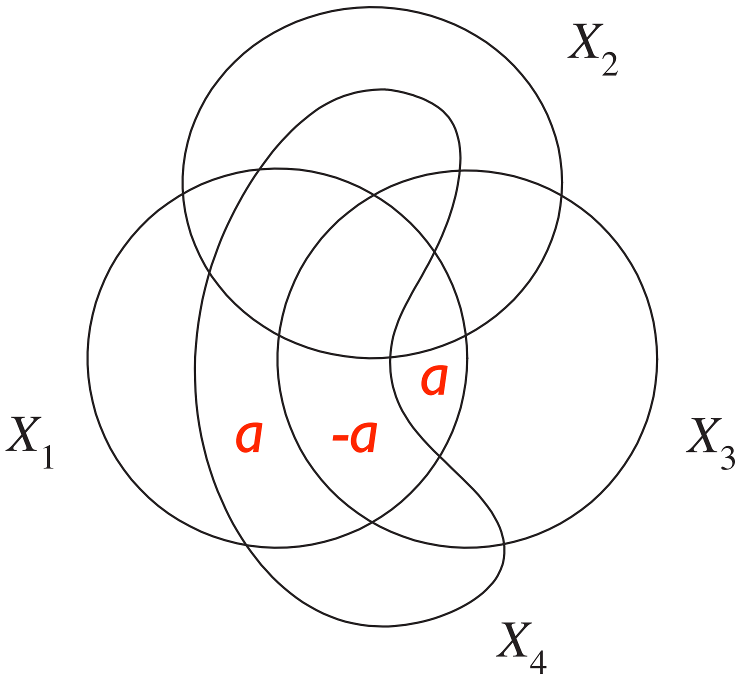
$$I(X_1; X_3; X_4|X_2) = -a.$$

2. The Markov subchain $X_1 \to X_2 \to X_4$ implies

$$0 = I(X_1; X_4|X_2) = I(X_1; X_3; X_4|X_2) + I(X_1; X_4|X_2, X_3).$$

Since $I(X_1; X_3; X_4|X_2) = -a$,

$$I(X_1; X_4|X_2, X_3) = a.$$



Structure of μ^* for $X_1 \to X_2 \to X_3 \to X_4$

1. The Markov subchain $X_1 \to X_2 \to X_3$ implies

$0 = I(X_1; X_3|X_2) = I(X_1; X_3; X_4|X_2) + I(X_1; X_3|X_2, X_4)$.

Let $I(X_1; X_3|X_2, X_4) = a \geq 0$. Then

$$I(X_1; X_3; X_4|X_2) = -a.$$

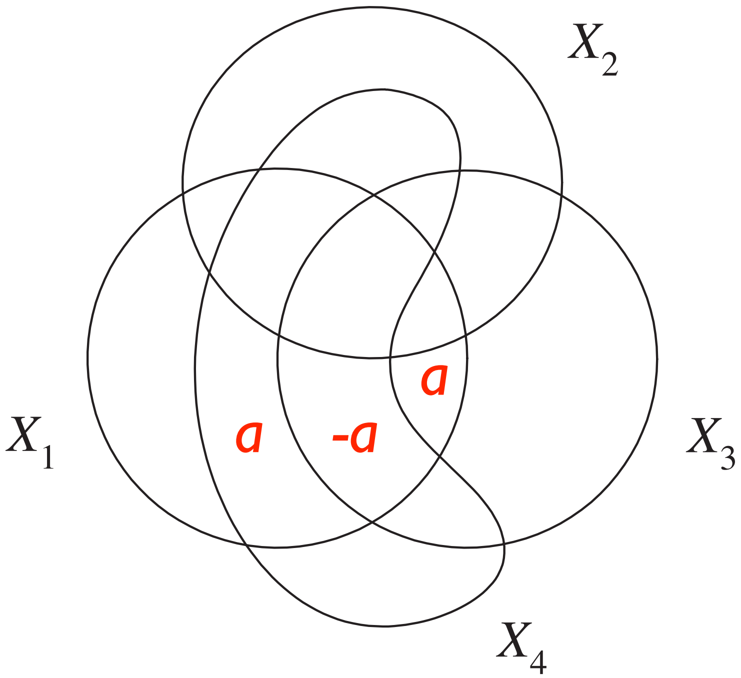
2. The Markov subchain $X_1 \to X_2 \to X_4$ implies

$0 = I(X_1; X_4|X_2) = I(X_1; X_3; X_4|X_2) + I(X_1; X_4|X_2, X_3)$.

Since $I(X_1; X_3; X_4|X_2) = -a$,

$$I(X_1; X_4|X_2, X_3) = a.$$

3. The Markov subchain $X_1 \to X_3 \to X_4$ implies



Structure of μ^* for $X_1 \to X_2 \to X_3 \to X_4$

1. The Markov subchain $X_1 \to X_2 \to X_3$ implies

$$0 = I(X_1; X_3|X_2) = I(X_1; X_3; X_4|X_2) + I(X_1; X_3|X_2, X_4).$$

Let $I(X_1; X_3|X_2, X_4) = a \geq 0$. Then

$$I(X_1; X_3; X_4|X_2) = -a.$$

2. The Markov subchain $X_1 \to X_2 \to X_4$ implies

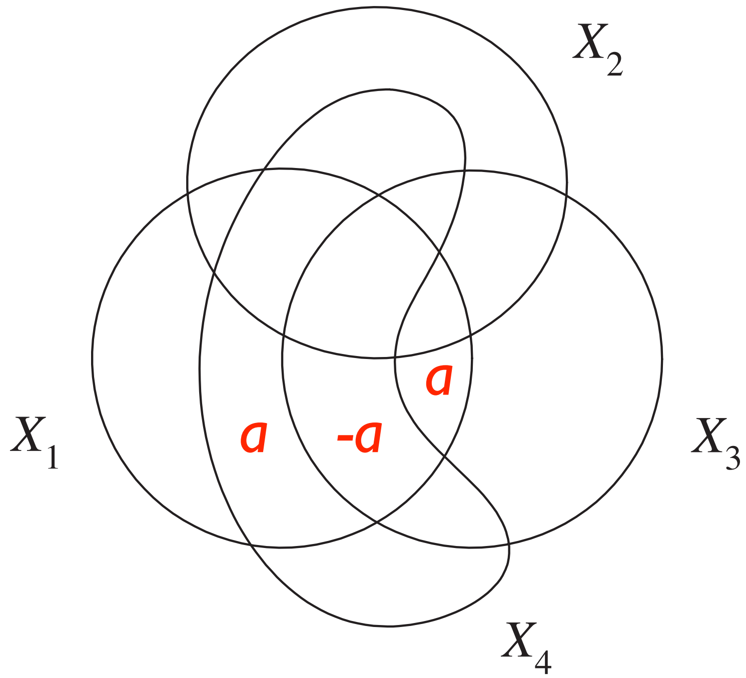
$$0 = I(X_1; X_4|X_2) = I(X_1; X_3; X_4|X_2) + I(X_1; X_4|X_2, X_3).$$

Since $I(X_1; X_3; X_4|X_2) = -a$,

$$I(X_1; X_4|X_2, X_3) = a.$$

3. The Markov subchain $X_1 \to X_3 \to X_4$ implies

$$0 = I(X_1; X_4|X_3) = I(X_1; X_2; X_4|X_3) + I(X_1; X_4|X_2, X_3).$$



Structure of μ^* for $X_1 \to X_2 \to X_3 \to X_4$

1. The Markov subchain $X_1 \to X_2 \to X_3$ implies

$0 = I(X_1; X_3|X_2) = I(X_1; X_3; X_4|X_2) + I(X_1; X_3|X_2, X_4)$.

Let $I(X_1; X_3|X_2, X_4) = a \geq 0$. Then

$$I(X_1; X_3; X_4|X_2) = -a.$$

2. The Markov subchain $X_1 \to X_2 \to X_4$ implies

$0 = I(X_1; X_4|X_2) = I(X_1; X_3; X_4|X_2) + I(X_1; X_4|X_2, X_3)$.

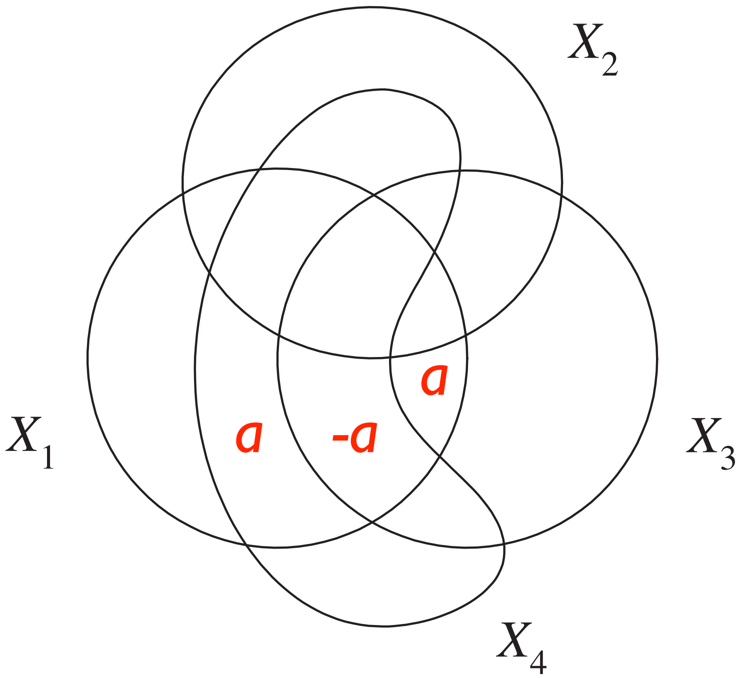
Since $I(X_1; X_3; X_4|X_2) = -a$,

$$I(X_1; X_4|X_2, X_3) = a.$$

3. The Markov subchain $X_1 \to X_3 \to X_4$ implies

$0 = I(X_1; X_4|X_3) = I(X_1; X_2; X_4|X_3) + I(X_1; X_4|X_2, X_3)$.

Since $I(X_1; X_4|X_2, X_3) = a$,



Structure of μ^* for $X_1 \to X_2 \to X_3 \to X_4$

1. The Markov subchain $X_1 \to X_2 \to X_3$ implies

$0 = I(X_1; X_3|X_2) = I(X_1; X_3; X_4|X_2) + I(X_1; X_3|X_2, X_4)$.

Let $I(X_1; X_3|X_2, X_4) = a \geq 0$. Then

$$I(X_1; X_3; X_4|X_2) = -a.$$

2. The Markov subchain $X_1 \to X_2 \to X_4$ implies

$0 = I(X_1; X_4|X_2) = I(X_1; X_3; X_4|X_2) + I(X_1; X_4|X_2, X_3)$.

Since $I(X_1; X_3; X_4|X_2) = -a$,

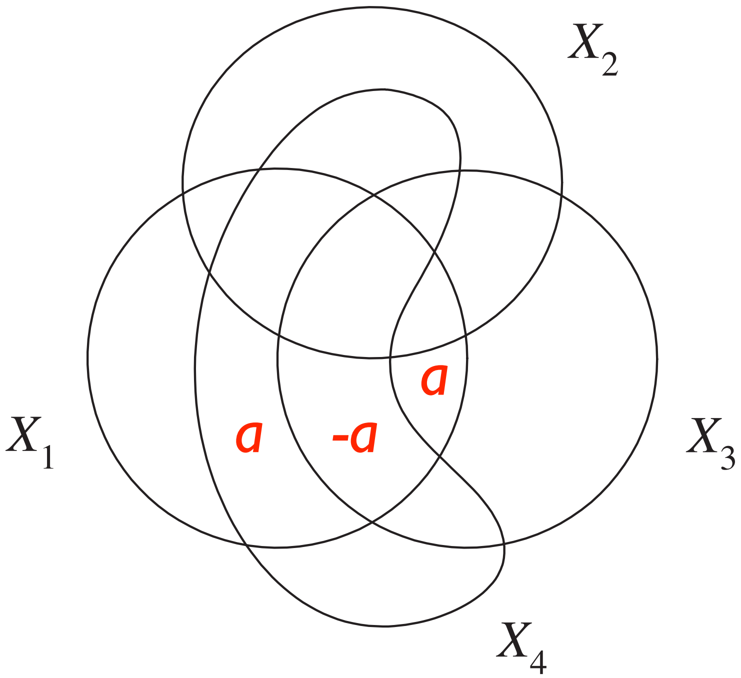
$$I(X_1; X_4|X_2, X_3) = a.$$

3. The Markov subchain $X_1 \to X_3 \to X_4$ implies

$0 = I(X_1; X_4|X_3) = I(X_1; X_2; X_4|X_3) + I(X_1; X_4|X_2, X_3)$.

Since $I(X_1; X_4|X_2, X_3) = a$,

$$I(X_1; X_2; X_4|X_3) = -a.$$



Structure of μ^* for $X_1 \to X_2 \to X_3 \to X_4$

1. The Markov subchain $X_1 \to X_2 \to X_3$ implies

$0 = I(X_1; X_3|X_2) = I(X_1; X_3; X_4|X_2) + I(X_1; X_3|X_2, X_4)$.

Let $I(X_1; X_3|X_2, X_4) = a \geq 0$. Then

$$I(X_1; X_3; X_4|X_2) = -a.$$

2. The Markov subchain $X_1 \to X_2 \to X_4$ implies

$0 = I(X_1; X_4|X_2) = I(X_1; X_3; X_4|X_2) + I(X_1; X_4|X_2, X_3)$.

Since $I(X_1; X_3; X_4|X_2) = -a$,

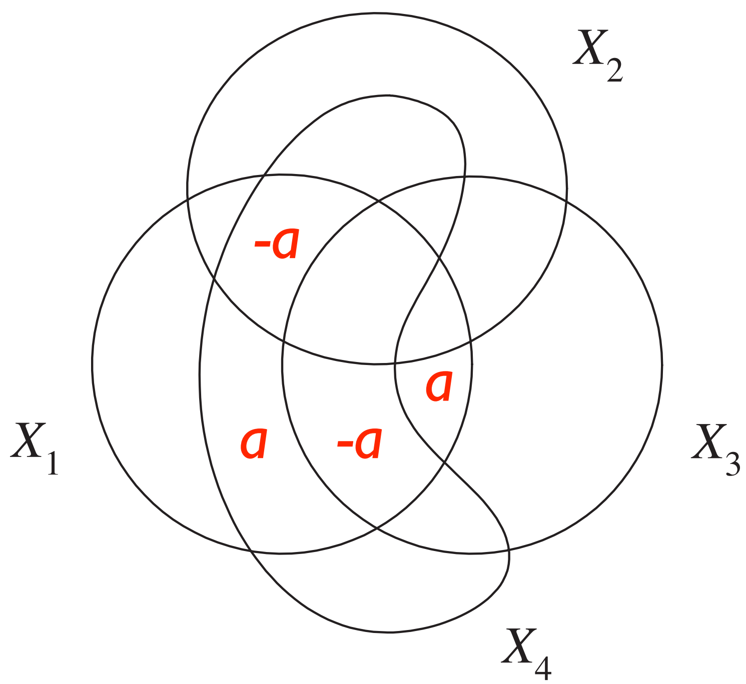
$$I(X_1; X_4|X_2, X_3) = a.$$

3. The Markov subchain $X_1 \to X_3 \to X_4$ implies

$0 = I(X_1; X_4|X_3) = I(X_1; X_2; X_4|X_3) + I(X_1; X_4|X_2, X_3)$.

Since $I(X_1; X_4|X_2, X_3) = a$,

$$I(X_1; X_2; X_4|X_3) = -a.$$



Structure of μ^* for $X_1 \to X_2 \to X_3 \to X_4$

1. The Markov subchain $X_1 \to X_2 \to X_3$ implies

$$0 = I(X_1; X_3|X_2) = I(X_1; X_3; X_4|X_2) + I(X_1; X_3|X_2, X_4).$$

Let $I(X_1; X_3|X_2, X_4) = a \geq 0$. Then

$$I(X_1; X_3; X_4|X_2) = -a.$$

2. The Markov subchain $X_1 \to X_2 \to X_4$ implies

$$0 = I(X_1; X_4|X_2) = I(X_1; X_3; X_4|X_2) + I(X_1; X_4|X_2, X_3).$$

Since $I(X_1; X_3; X_4|X_2) = -a$,

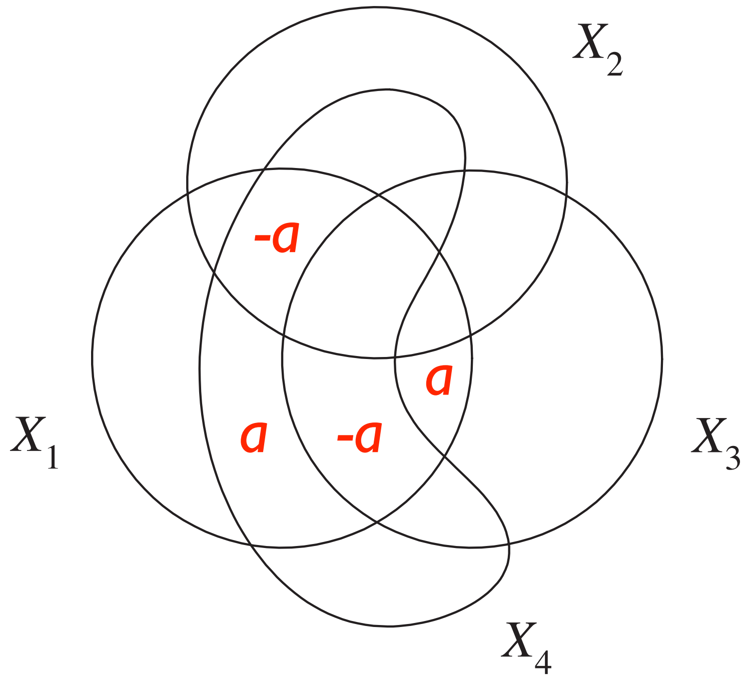
$$I(X_1; X_4|X_2, X_3) = a.$$

3. The Markov subchain $X_1 \to X_3 \to X_4$ implies

$$0 = I(X_1; X_4|X_3) = I(X_1; X_2; X_4|X_3) + I(X_1; X_4|X_2, X_3).$$

Since $I(X_1; X_4|X_2, X_3) = a$,

$$I(X_1; X_2; X_4|X_3) = -a.$$



Structure of μ^* for $X_1 \to X_2 \to X_3 \to X_4$

1. The Markov subchain $X_1 \to X_2 \to X_3$ implies

$0 = I(X_1; X_3|X_2) = I(X_1; X_3; X_4|X_2) + I(X_1; X_3|X_2, X_4)$.

Let $I(X_1; X_3|X_2, X_4) = a \geq 0$. Then

$$I(X_1; X_3; X_4|X_2) = -a.$$

2. The Markov subchain $X_1 \to X_2 \to X_4$ implies

$0 = I(X_1; X_4|X_2) = I(X_1; X_3; X_4|X_2) + I(X_1; X_4|X_2, X_3)$.

Since $I(X_1; X_3; X_4|X_2) = -a$,

$$I(X_1; X_4|X_2, X_3) = a.$$

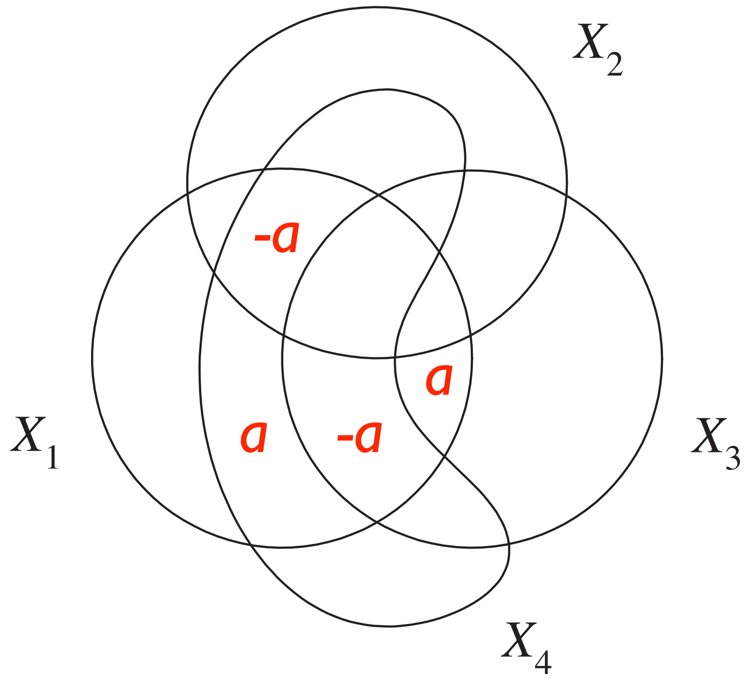
3. The Markov subchain $X_1 \to X_3 \to X_4$ implies

$0 = I(X_1; X_4|X_3) = I(X_1; X_2; X_4|X_3) + I(X_1; X_4|X_2, X_3)$.

Since $I(X_1; X_4|X_2, X_3) = a$,

$$I(X_1; X_2; X_4|X_3) = -a.$$

4. The Markov subchain $X_2 \to X_3 \to X_4$ implies



Structure of μ^* for $X_1 \to X_2 \to X_3 \to X_4$

1. The Markov subchain $X_1 \to X_2 \to X_3$ implies

$0 = I(X_1; X_3|X_2) = I(X_1; X_3; X_4|X_2) + I(X_1; X_3|X_2, X_4)$.

Let $I(X_1; X_3|X_2, X_4) = a \geq 0$. Then

$$I(X_1; X_3; X_4|X_2) = -a.$$

2. The Markov subchain $X_1 \to X_2 \to X_4$ implies

$0 = I(X_1; X_4|X_2) = I(X_1; X_3; X_4|X_2) + I(X_1; X_4|X_2, X_3)$.

Since $I(X_1; X_3; X_4|X_2) = -a$,

$$I(X_1; X_4|X_2, X_3) = a.$$

3. The Markov subchain $X_1 \to X_3 \to X_4$ implies

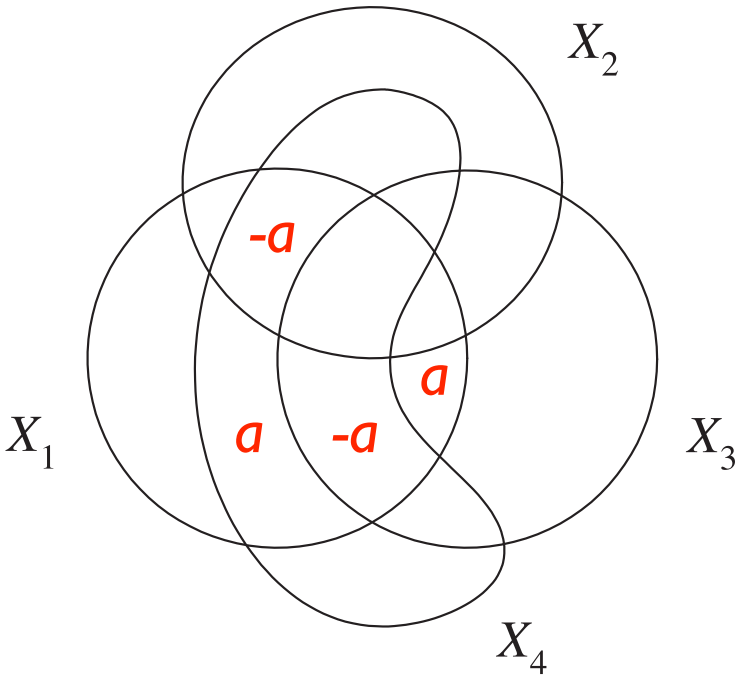
$0 = I(X_1; X_4|X_3) = I(X_1; X_2; X_4|X_3) + I(X_1; X_4|X_2, X_3)$.

Since $I(X_1; X_4|X_2, X_3) = a$,

$$I(X_1; X_2; X_4|X_3) = -a.$$

4. The Markov subchain $X_2 \to X_3 \to X_4$ implies

$0 = I(X_2; X_4|X_3) = I(X_1; X_2; X_4|X_3) + I(X_2; X_4|X_1, X_3)$.



Structure of μ^* for $X_1 \to X_2 \to X_3 \to X_4$

1. The Markov subchain $X_1 \to X_2 \to X_3$ implies

$0 = I(X_1; X_3|X_2) = I(X_1; X_3; X_4|X_2) + I(X_1; X_3|X_2, X_4)$.

Let $I(X_1; X_3|X_2, X_4) = a \geq 0$. Then

$$I(X_1; X_3; X_4|X_2) = -a.$$

2. The Markov subchain $X_1 \to X_2 \to X_4$ implies

$0 = I(X_1; X_4|X_2) = I(X_1; X_3; X_4|X_2) + I(X_1; X_4|X_2, X_3)$.

Since $I(X_1; X_3; X_4|X_2) = -a$,

$$I(X_1; X_4|X_2, X_3) = a.$$

3. The Markov subchain $X_1 \to X_3 \to X_4$ implies

$0 = I(X_1; X_4|X_3) = I(X_1; X_2; X_4|X_3) + I(X_1; X_4|X_2, X_3)$.

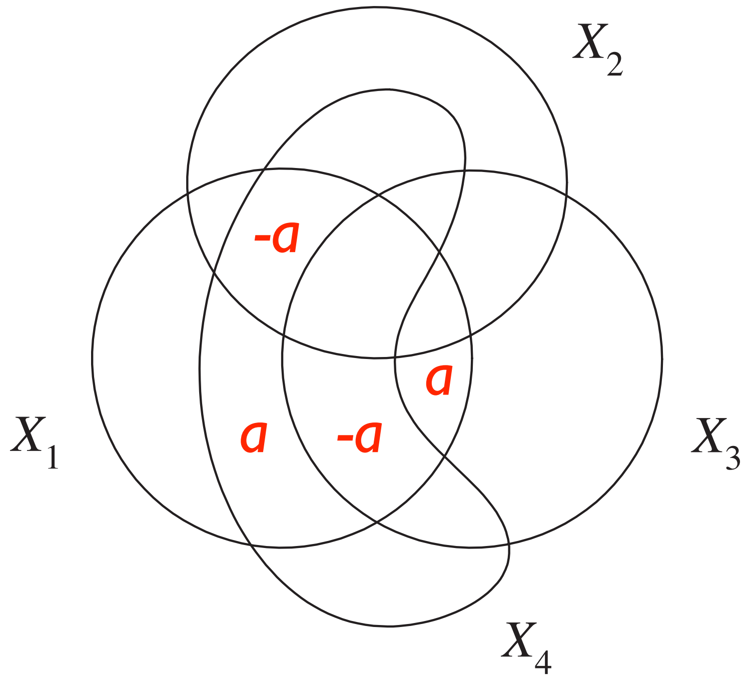
Since $I(X_1; X_4|X_2, X_3) = a$,

$$I(X_1; X_2; X_4|X_3) = -a.$$

4. The Markov subchain $X_2 \to X_3 \to X_4$ implies

$0 = I(X_2; X_4|X_3) = I(X_1; X_2; X_4|X_3) + I(X_2; X_4|X_1, X_3)$.

Since $I(X_1; X_2; X_4|X_3) = -a$,



Structure of μ^* for $X_1 \to X_2 \to X_3 \to X_4$

1. The Markov subchain $X_1 \to X_2 \to X_3$ implies

$0 = I(X_1; X_3|X_2) = I(X_1; X_3; X_4|X_2) + I(X_1; X_3|X_2, X_4)$.

Let $I(X_1; X_3|X_2, X_4) = a \geq 0$. Then

$$I(X_1; X_3; X_4|X_2) = -a.$$

2. The Markov subchain $X_1 \to X_2 \to X_4$ implies

$0 = I(X_1; X_4|X_2) = I(X_1; X_3; X_4|X_2) + I(X_1; X_4|X_2, X_3)$.

Since $I(X_1; X_3; X_4|X_2) = -a$,

$$I(X_1; X_4|X_2, X_3) = a.$$

3. The Markov subchain $X_1 \to X_3 \to X_4$ implies

$0 = I(X_1; X_4|X_3) = I(X_1; X_2; X_4|X_3) + I(X_1; X_4|X_2, X_3)$.

Since $I(X_1; X_4|X_2, X_3) = a$,

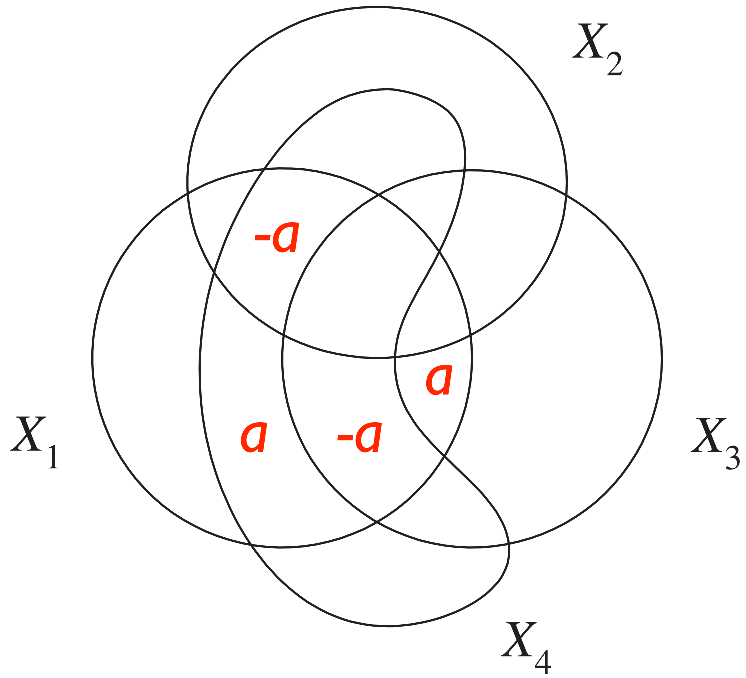
$$I(X_1; X_2; X_4|X_3) = -a.$$

4. The Markov subchain $X_2 \to X_3 \to X_4$ implies

$0 = I(X_2; X_4|X_3) = I(X_1; X_2; X_4|X_3) + I(X_2; X_4|X_1, X_3)$.

Since $I(X_1; X_2; X_4|X_3) = -a$,

$$I(X_2; X_4|X_1, X_3) = a.$$



Structure of μ^* for $X_1 \to X_2 \to X_3 \to X_4$

1. The Markov subchain $X_1 \to X_2 \to X_3$ implies

$$0 = I(X_1; X_3|X_2) = I(X_1; X_3; X_4|X_2) + I(X_1; X_3|X_2, X_4).$$

Let $I(X_1; X_3|X_2, X_4) = a \geq 0$. Then

$$I(X_1; X_3; X_4|X_2) = -a.$$

2. The Markov subchain $X_1 \to X_2 \to X_4$ implies

$$0 = I(X_1; X_4|X_2) = I(X_1; X_3; X_4|X_2) + I(X_1; X_4|X_2, X_3).$$

Since $I(X_1; X_3; X_4|X_2) = -a$,

$$I(X_1; X_4|X_2, X_3) = a.$$

3. The Markov subchain $X_1 \to X_3 \to X_4$ implies

$$0 = I(X_1; X_4|X_3) = I(X_1; X_2; X_4|X_3) + I(X_1; X_4|X_2, X_3).$$

Since $I(X_1; X_4|X_2, X_3) = a$,

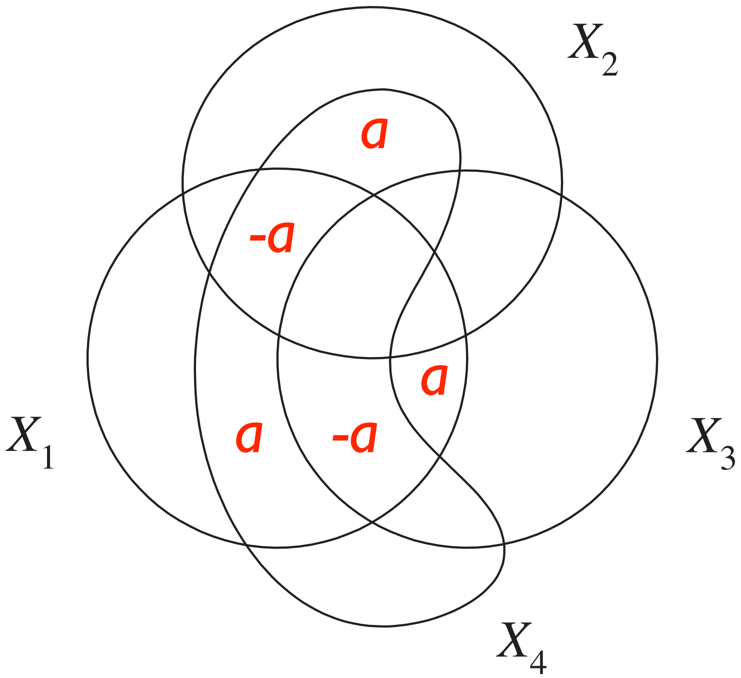
$$I(X_1; X_2; X_4|X_3) = -a.$$

4. The Markov subchain $X_2 \to X_3 \to X_4$ implies

$$0 = I(X_2; X_4|X_3) = \underline{I(X_1; X_2; X_4|X_3)} + \underline{I(X_2; X_4|X_1, X_3)}.$$

Since $\underline{I(X_1; X_2; X_4|X_3) = -a}$,

$$\underline{I(X_2; X_4|X_1, X_3) = a.}$$



Structure of μ^* for $X_1 \to X_2 \to X_3 \to X_4$

1. The Markov subchain $X_1 \to X_2 \to X_3$ implies

$$0 = I(X_1; X_3|X_2) = I(X_1; X_3; X_4|X_2) + I(X_1; X_3|X_2, X_4).$$

Let $I(X_1; X_3|X_2, X_4) = a \geq 0$. Then

$$I(X_1; X_3; X_4|X_2) = -a.$$

2. The Markov subchain $X_1 \to X_2 \to X_4$ implies

$$0 = I(X_1; X_4|X_2) = I(X_1; X_3; X_4|X_2) + I(X_1; X_4|X_2, X_3).$$

Since $I(X_1; X_3; X_4|X_2) = -a$,

$$I(X_1; X_4|X_2, X_3) = a.$$

3. The Markov subchain $X_1 \to X_3 \to X_4$ implies

$$0 = I(X_1; X_4|X_3) = I(X_1; X_2; X_4|X_3) + I(X_1; X_4|X_2, X_3).$$

Since $I(X_1; X_4|X_2, X_3) = a$,

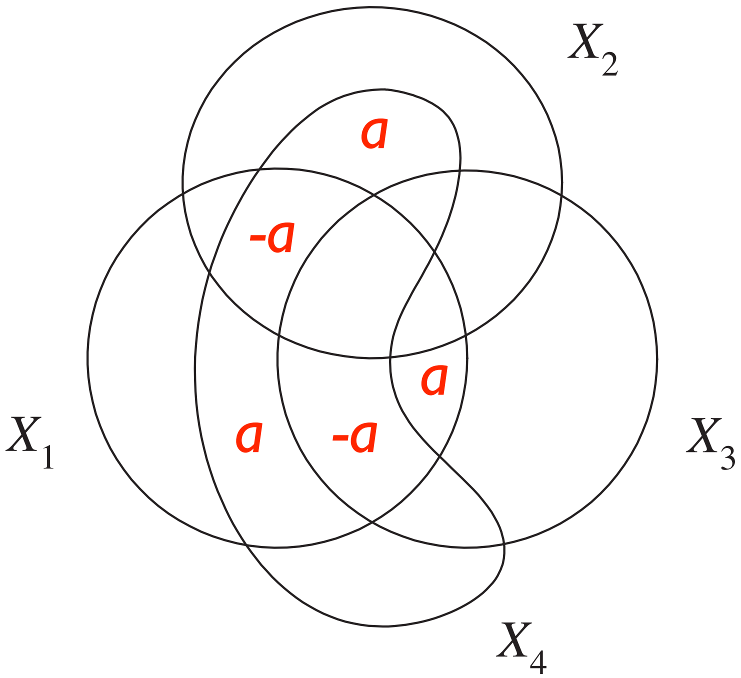
$$I(X_1; X_2; X_4|X_3) = -a.$$

4. The Markov subchain $X_2 \to X_3 \to X_4$ implies

$$0 = I(X_2; X_4|X_3) = I(X_1; X_2; X_4|X_3) + I(X_2; X_4|X_1, X_3).$$

Since $I(X_1; X_2; X_4|X_3) = -a$,

$$I(X_2; X_4|X_1, X_3) = a.$$



Structure of μ^* for $X_1 \to X_2 \to X_3 \to X_4$

1. The Markov subchain $X_1 \to X_2 \to X_3$ implies

$0 = I(X_1; X_3|X_2) = I(X_1; X_3; X_4|X_2) + I(X_1; X_3|X_2, X_4)$.

Let $I(X_1; X_3|X_2, X_4) = a \geq 0$. Then

$$I(X_1; X_3; X_4|X_2) = -a.$$

2. The Markov subchain $X_1 \to X_2 \to X_4$ implies

$0 = I(X_1; X_4|X_2) = I(X_1; X_3; X_4|X_2) + I(X_1; X_4|X_2, X_3)$.

Since $I(X_1; X_3; X_4|X_2) = -a$,

$$I(X_1; X_4|X_2, X_3) = a.$$

3. The Markov subchain $X_1 \to X_3 \to X_4$ implies

$0 = I(X_1; X_4|X_3) = I(X_1; X_2; X_4|X_3) + I(X_1; X_4|X_2, X_3)$.

Since $I(X_1; X_4|X_2, X_3) = a$,

$$I(X_1; X_2; X_4|X_3) = -a.$$

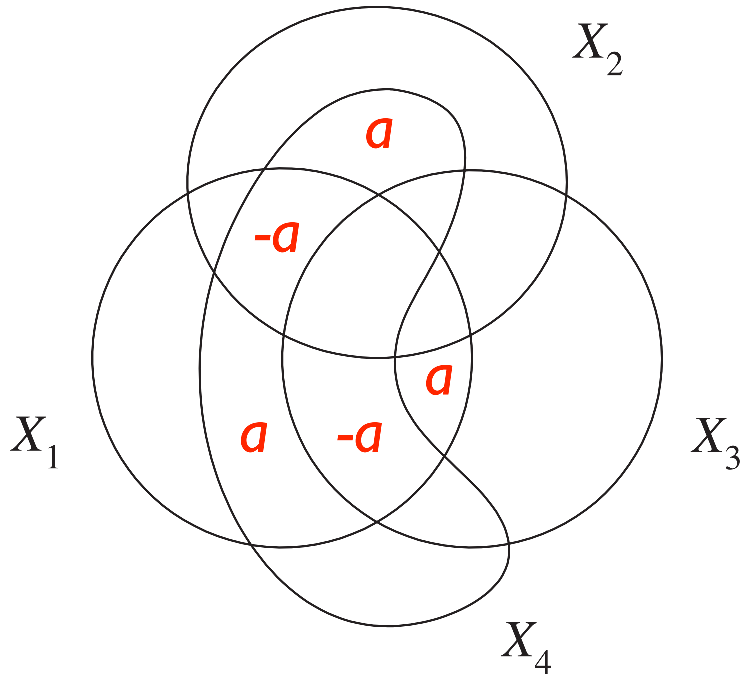
4. The Markov subchain $X_2 \to X_3 \to X_4$ implies

$0 = I(X_2; X_4|X_3) = I(X_1; X_2; X_4|X_3) + I(X_2; X_4|X_1, X_3)$.

Since $I(X_1; X_2; X_4|X_3) = -a$,

$$I(X_2; X_4|X_1, X_3) = a.$$

5. The Markov subchain $(X_1, X_2) \to X_3 \to X_4$ implies



Structure of μ^* for $X_1 \to X_2 \to X_3 \to X_4$

1. The Markov subchain $X_1 \to X_2 \to X_3$ implies

$$0 = I(X_1; X_3|X_2) = I(X_1; X_3; X_4|X_2) + I(X_1; X_3|X_2, X_4).$$

Let $I(X_1; X_3|X_2, X_4) = a \geq 0$. Then

$$I(X_1; X_3; X_4|X_2) = -a.$$

2. The Markov subchain $X_1 \to X_2 \to X_4$ implies

$$0 = I(X_1; X_4|X_2) = I(X_1; X_3; X_4|X_2) + I(X_1; X_4|X_2, X_3).$$

Since $I(X_1; X_3; X_4|X_2) = -a$,

$$I(X_1; X_4|X_2, X_3) = a.$$

3. The Markov subchain $X_1 \to X_3 \to X_4$ implies

$$0 = I(X_1; X_4|X_3) = I(X_1; X_2; X_4|X_3) + I(X_1; X_4|X_2, X_3).$$

Since $I(X_1; X_4|X_2, X_3) = a$,

$$I(X_1; X_2; X_4|X_3) = -a.$$

4. The Markov subchain $X_2 \to X_3 \to X_4$ implies

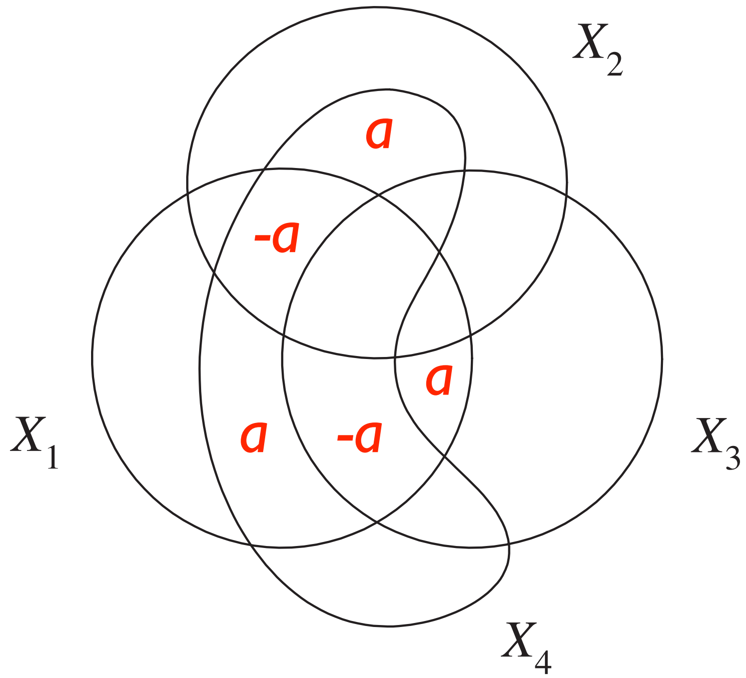
$$0 = I(X_2; X_4|X_3) = I(X_1; X_2; X_4|X_3) + I(X_2; X_4|X_1, X_3).$$

Since $I(X_1; X_2; X_4|X_3) = -a$,

$$I(X_2; X_4|X_1, X_3) = a.$$

5. The Markov subchain $(X_1, X_2) \to X_3 \to X_4$ implies

$$0 = I(X_1, X_2; X_4|X_3) =$$



Structure of μ^* for $X_1 \to X_2 \to X_3 \to X_4$

1. The Markov subchain $X_1 \to X_2 \to X_3$ implies

$$0 = I(X_1; X_3|X_2) = I(X_1; X_3; X_4|X_2) + I(X_1; X_3|X_2, X_4).$$

Let $I(X_1; X_3|X_2, X_4) = a \geq 0$. Then

$$I(X_1; X_3; X_4|X_2) = -a.$$

2. The Markov subchain $X_1 \to X_2 \to X_4$ implies

$$0 = I(X_1; X_4|X_2) = I(X_1; X_3; X_4|X_2) + I(X_1; X_4|X_2, X_3).$$

Since $I(X_1; X_3; X_4|X_2) = -a$,

$$I(X_1; X_4|X_2, X_3) = a.$$

3. The Markov subchain $X_1 \to X_3 \to X_4$ implies

$$0 = I(X_1; X_4|X_3) = I(X_1; X_2; X_4|X_3) + I(X_1; X_4|X_2, X_3).$$

Since $I(X_1; X_4|X_2, X_3) = a$,

$$I(X_1; X_2; X_4|X_3) = -a.$$

4. The Markov subchain $X_2 \to X_3 \to X_4$ implies

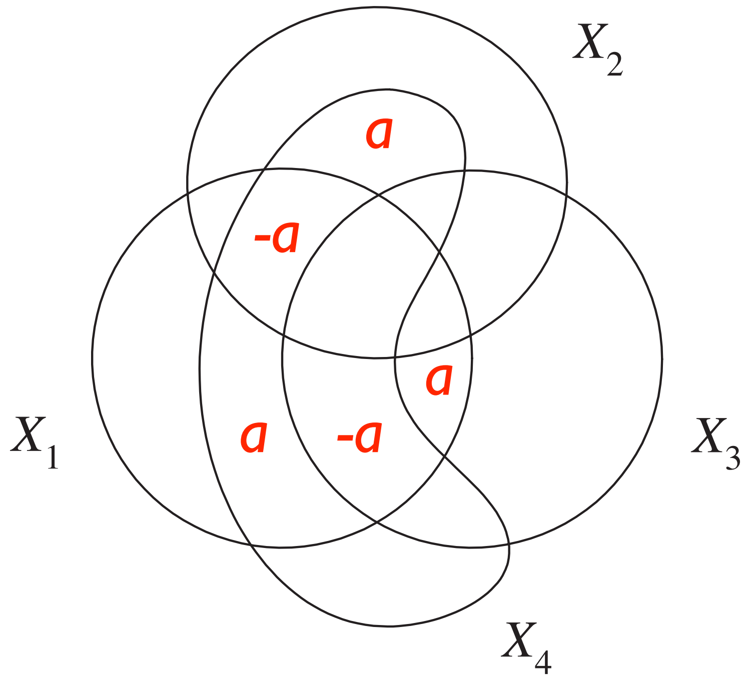
$$0 = I(X_2; X_4|X_3) = I(X_1; X_2; X_4|X_3) + I(X_2; X_4|X_1, X_3).$$

Since $I(X_1; X_2; X_4|X_3) = -a$,

$$I(X_2; X_4|X_1, X_3) = a.$$

5. The Markov subchain $(X_1, X_2) \to X_3 \to X_4$ implies

$$0 = I(X_1, X_2; X_4|X_3) = I(X_1; X_4|X_2, X_3) + I(X_1; X_2; X_4|X_3) + I(X_2; X_4|X_1, X_3).$$



Structure of μ^* for $X_1 \to X_2 \to X_3 \to X_4$

1. The Markov subchain $X_1 \to X_2 \to X_3$ implies

$$0 = I(X_1; X_3|X_2) = I(X_1; X_3; X_4|X_2) + I(X_1; X_3|X_2, X_4).$$

Let $I(X_1; X_3|X_2, X_4) = a \geq 0$. Then

$$I(X_1; X_3; X_4|X_2) = -a.$$

2. The Markov subchain $X_1 \to X_2 \to X_4$ implies

$$0 = I(X_1; X_4|X_2) = I(X_1; X_3; X_4|X_2) + I(X_1; X_4|X_2, X_3).$$

Since $I(X_1; X_3; X_4|X_2) = -a$,

$$I(X_1; X_4|X_2, X_3) = a.$$

3. The Markov subchain $X_1 \to X_3 \to X_4$ implies

$$0 = I(X_1; X_4|X_3) = I(X_1; X_2; X_4|X_3) + I(X_1; X_4|X_2, X_3).$$

Since $I(X_1; X_4|X_2, X_3) = a$,

$$I(X_1; X_2; X_4|X_3) = -a.$$

4. The Markov subchain $X_2 \to X_3 \to X_4$ implies

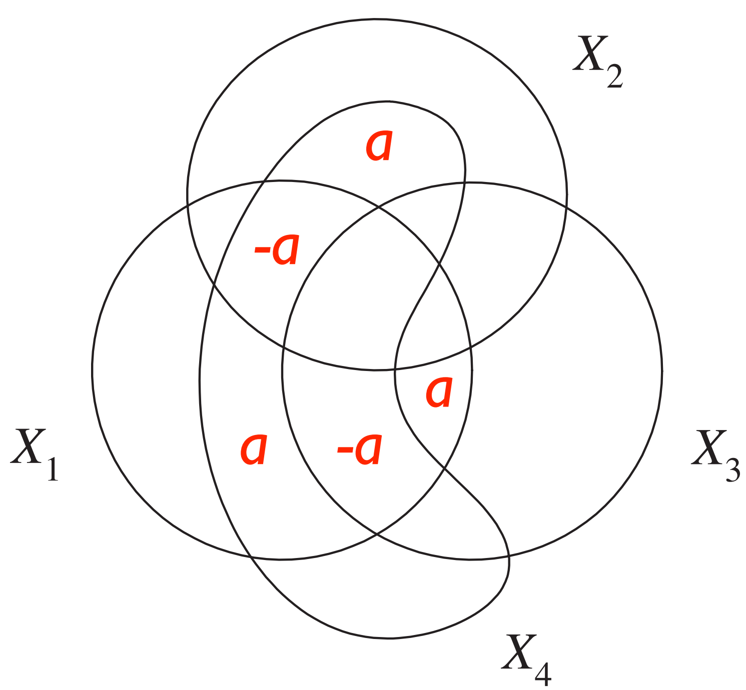
$$0 = I(X_2; X_4|X_3) = I(X_1; X_2; X_4|X_3) + I(X_2; X_4|X_1, X_3).$$

Since $I(X_1; X_2; X_4|X_3) = -a$,

$$I(X_2; X_4|X_1, X_3) = a.$$

5. The Markov subchain $(X_1, X_2) \to X_3 \to X_4$ implies

$$0 = I(X_1, X_2; X_4|X_3) = I(X_1; X_4|X_2, X_3) + I(X_1; X_2; X_4|X_3) + I(X_2; X_4|X_1, X_3).$$



Structure of μ^* for $X_1 \to X_2 \to X_3 \to X_4$

1. The Markov subchain $X_1 \to X_2 \to X_3$ implies

$$0 = I(X_1; X_3|X_2) = I(X_1; X_3; X_4|X_2) + I(X_1; X_3|X_2, X_4).$$

Let $I(X_1; X_3|X_2, X_4) = a \geq 0$. Then

$$I(X_1; X_3; X_4|X_2) = -a.$$

2. The Markov subchain $X_1 \to X_2 \to X_4$ implies

$$0 = I(X_1; X_4|X_2) = I(X_1; X_3; X_4|X_2) + I(X_1; X_4|X_2, X_3).$$

Since $I(X_1; X_3; X_4|X_2) = -a$,

$$I(X_1; X_4|X_2, X_3) = a.$$

3. The Markov subchain $X_1 \to X_3 \to X_4$ implies

$$0 = I(X_1; X_4|X_3) = I(X_1; X_2; X_4|X_3) + I(X_1; X_4|X_2, X_3).$$

Since $I(X_1; X_4|X_2, X_3) = a$,

$$I(X_1; X_2; X_4|X_3) = -a.$$

4. The Markov subchain $X_2 \to X_3 \to X_4$ implies

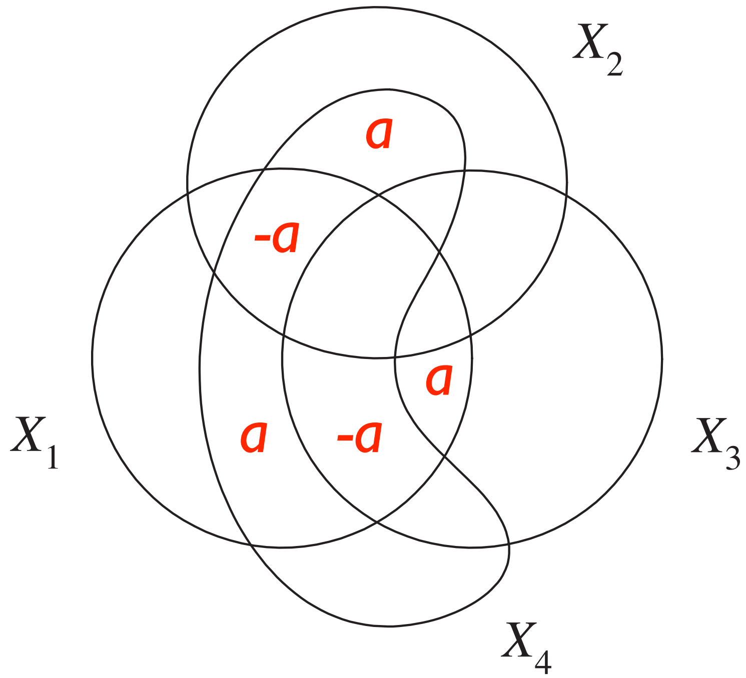
$$0 = I(X_2; X_4|X_3) = I(X_1; X_2; X_4|X_3) + I(X_2; X_4|X_1, X_3).$$

Since $I(X_1; X_2; X_4|X_3) = -a$,

$$I(X_2; X_4|X_1, X_3) = a.$$

5. The Markov subchain $(X_1, X_2) \to X_3 \to X_4$ implies

$$0 = I(X_1, X_2; X_4|X_3) = I(X_1; X_4|X_2, X_3) + I(X_1; X_2; X_4|X_3) + I(X_2; X_4|X_1, X_3).$$



Structure of μ^* for $X_1 \to X_2 \to X_3 \to X_4$

1. The Markov subchain $X_1 \to X_2 \to X_3$ implies

$$0 = I(X_1; X_3|X_2) = I(X_1; X_3; X_4|X_2) + I(X_1; X_3|X_2, X_4).$$

Let $I(X_1; X_3|X_2, X_4) = a \geq 0$. Then

$$I(X_1; X_3; X_4|X_2) = -a.$$

2. The Markov subchain $X_1 \to X_2 \to X_4$ implies

$$0 = I(X_1; X_4|X_2) = I(X_1; X_3; X_4|X_2) + I(X_1; X_4|X_2, X_3).$$

Since $I(X_1; X_3; X_4|X_2) = -a$,

$$I(X_1; X_4|X_2, X_3) = a.$$

3. The Markov subchain $X_1 \to X_3 \to X_4$ implies

$$0 = I(X_1; X_4|X_3) = I(X_1; X_2; X_4|X_3) + I(X_1; X_4|X_2, X_3).$$

Since $I(X_1; X_4|X_2, X_3) = a$,

$$I(X_1; X_2; X_4|X_3) = -a.$$

4. The Markov subchain $X_2 \to X_3 \to X_4$ implies

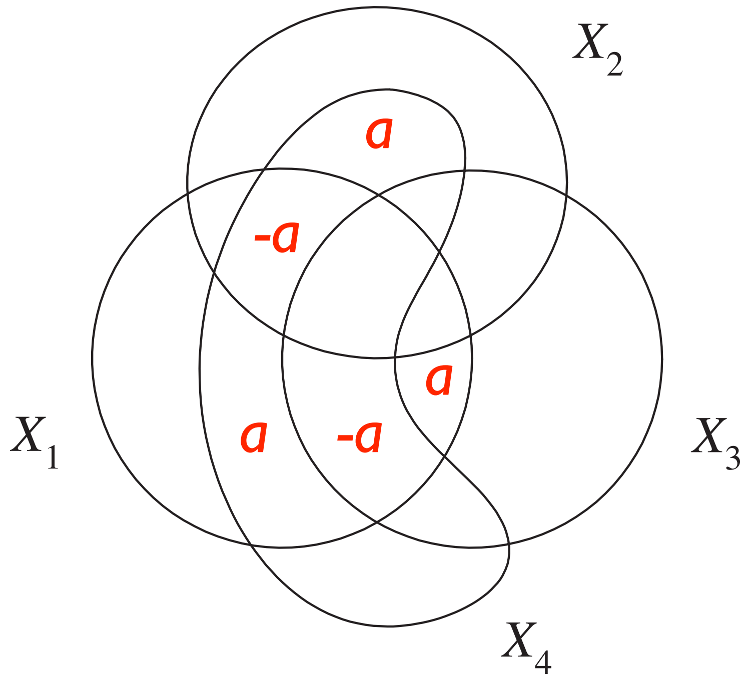
$$0 = I(X_2; X_4|X_3) = I(X_1; X_2; X_4|X_3) + I(X_2; X_4|X_1, X_3).$$

Since $I(X_1; X_2; X_4|X_3) = -a$,

$$I(X_2; X_4|X_1, X_3) = a.$$

5. The Markov subchain $(X_1, X_2) \to X_3 \to X_4$ implies

$$0 = I(X_1, X_2; X_4|X_3) = I(X_1; X_4|X_2, X_3) + I(X_1; X_2; X_4|X_3) + I(X_2; X_4|X_1, X_3).$$



Structure of μ^* for $X_1 \to X_2 \to X_3 \to X_4$

1. The Markov subchain $X_1 \to X_2 \to X_3$ implies

$$0 = I(X_1; X_3|X_2) = I(X_1; X_3; X_4|X_2) + I(X_1; X_3|X_2, X_4).$$

Let $I(X_1; X_3|X_2, X_4) = a \geq 0$. Then

$$I(X_1; X_3; X_4|X_2) = -a.$$

2. The Markov subchain $X_1 \to X_2 \to X_4$ implies

$$0 = I(X_1; X_4|X_2) = I(X_1; X_3; X_4|X_2) + I(X_1; X_4|X_2, X_3).$$

Since $I(X_1; X_3; X_4|X_2) = -a$,

$$I(X_1; X_4|X_2, X_3) = a.$$

3. The Markov subchain $X_1 \to X_3 \to X_4$ implies

$$0 = I(X_1; X_4|X_3) = I(X_1; X_2; X_4|X_3) + I(X_1; X_4|X_2, X_3).$$

Since $I(X_1; X_4|X_2, X_3) = a$,

$$I(X_1; X_2; X_4|X_3) = -a.$$

4. The Markov subchain $X_2 \to X_3 \to X_4$ implies

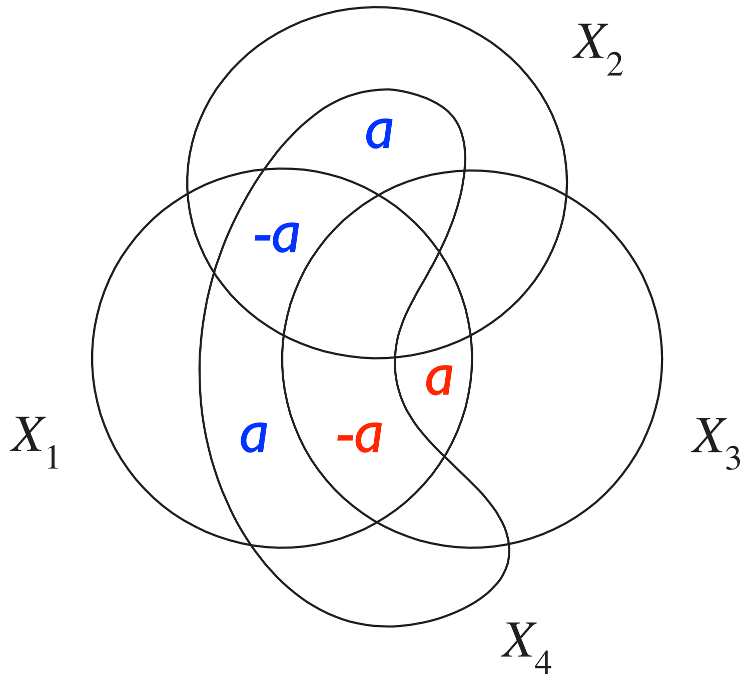
$$0 = I(X_2; X_4|X_3) = I(X_1; X_2; X_4|X_3) + I(X_2; X_4|X_1, X_3).$$

Since $I(X_1; X_2; X_4|X_3) = -a$,

$$I(X_2; X_4|X_1, X_3) = a.$$

5. The Markov subchain $(X_1, X_2) \to X_3 \to X_4$ implies

$$0 = I(X_1, X_2; X_4|X_3) = I(X_1; X_4|X_2, X_3) + I(X_1; X_2; X_4|X_3) + I(X_2; X_4|X_1, X_3).$$



Structure of μ^* for $X_1 \to X_2 \to X_3 \to X_4$

1. The Markov subchain $X_1 \to X_2 \to X_3$ implies

$0 = I(X_1;X_3|X_2) = I(X_1;X_3;X_4|X_2) + I(X_1;X_3|X_2,X_4)$.

Let $I(X_1;X_3|X_2,X_4) = a \geq 0$. Then

$$I(X_1;X_3;X_4|X_2) = -a.$$

2. The Markov subchain $X_1 \to X_2 \to X_4$ implies

$0 = I(X_1;X_4|X_2) = I(X_1;X_3;X_4|X_2) + I(X_1;X_4|X_2,X_3)$.

Since $I(X_1;X_3;X_4|X_2) = -a$,

$$I(X_1;X_4|X_2,X_3) = a.$$

3. The Markov subchain $X_1 \to X_3 \to X_4$ implies

$0 = I(X_1;X_4|X_3) = I(X_1;X_2;X_4|X_3) + I(X_1;X_4|X_2,X_3)$.

Since $I(X_1;X_4|X_2,X_3) = a$,

$$I(X_1;X_2;X_4|X_3) = -a.$$

4. The Markov subchain $X_2 \to X_3 \to X_4$ implies

$0 = I(X_2;X_4|X_3) = I(X_1;X_2;X_4|X_3) + I(X_2;X_4|X_1,X_3)$.

Since $I(X_1;X_2;X_4|X_3) = -a$,

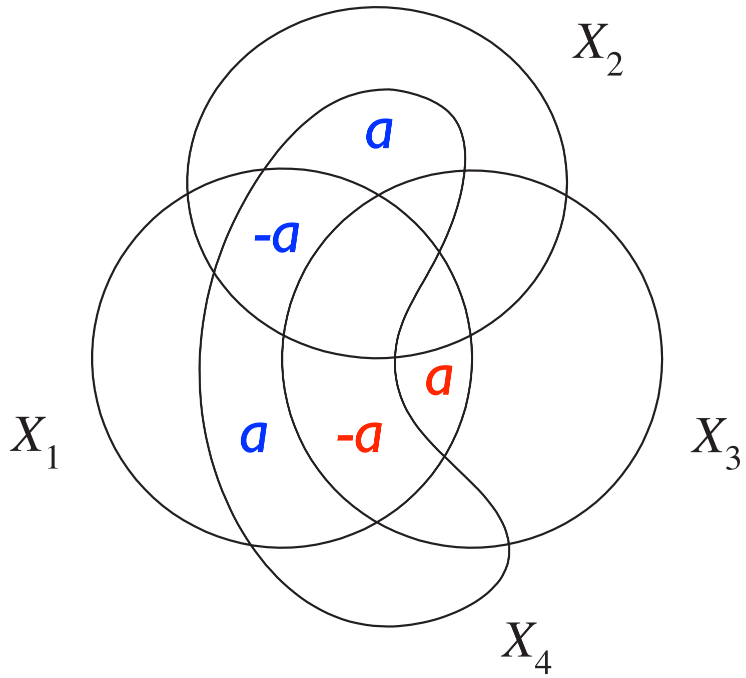
$$I(X_2;X_4|X_1,X_3) = a.$$

5. The Markov subchain $(X_1,X_2) \to X_3 \to X_4$ implies

$$0 = I(X_1,X_2;X_4|X_3) = I(X_1;X_4|X_2,X_3) + I(X_1;X_2;X_4|X_3) + I(X_2;X_4|X_1,X_3).$$

Then

$$0 = a - a + a = a.$$



Structure of μ^* for $X_1 \to X_2 \to X_3 \to X_4$

1. The Markov subchain $X_1 \to X_2 \to X_3$ implies

$$0 = I(X_1;X_3|X_2) = I(X_1;X_3;X_4|X_2) + I(X_1;X_3|X_2,X_4).$$

Let $I(X_1;X_3|X_2,X_4) = a \geq 0$. Then

$$I(X_1;X_3;X_4|X_2) = -a.$$

2. The Markov subchain $X_1 \to X_2 \to X_4$ implies

$$0 = I(X_1;X_4|X_2) = I(X_1;X_3;X_4|X_2) + I(X_1;X_4|X_2,X_3).$$

Since $I(X_1;X_3;X_4|X_2) = -a$,

$$I(X_1;X_4|X_2,X_3) = a.$$

3. The Markov subchain $X_1 \to X_3 \to X_4$ implies

$$0 = I(X_1;X_4|X_3) = I(X_1;X_2;X_4|X_3) + I(X_1;X_4|X_2,X_3).$$

Since $I(X_1;X_4|X_2,X_3) = a$,

$$I(X_1;X_2;X_4|X_3) = -a.$$

4. The Markov subchain $X_2 \to X_3 \to X_4$ implies

$$0 = I(X_2;X_4|X_3) = I(X_1;X_2;X_4|X_3) + I(X_2;X_4|X_1,X_3).$$

Since $I(X_1;X_2;X_4|X_3) = -a$,

$$I(X_2;X_4|X_1,X_3) = a.$$

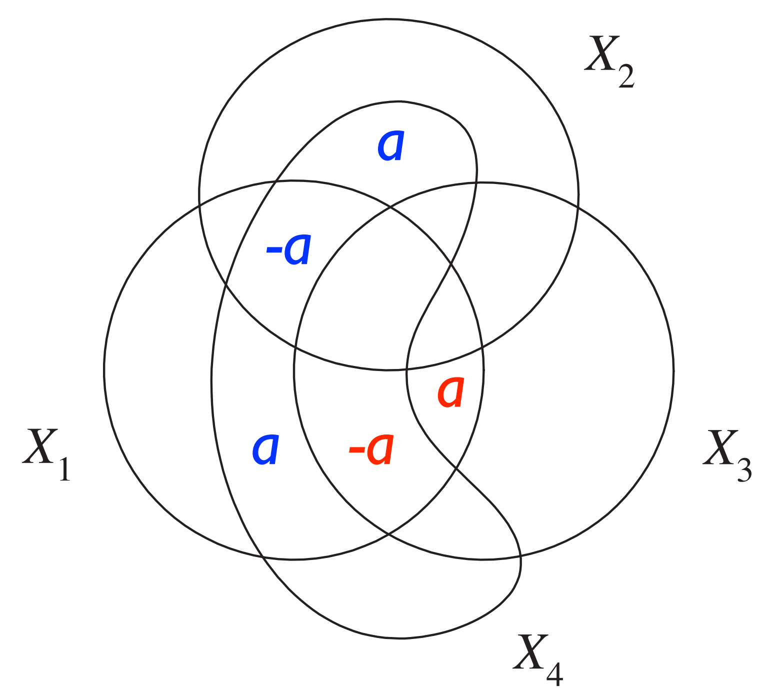
5. The Markov subchain $(X_1,X_2) \to X_3 \to X_4$ implies

$$0 = I(X_1,X_2;X_4|X_3) = I(X_1;X_4|X_2,X_3) + I(X_1;X_2;X_4|X_3) + I(X_2;X_4|X_1,X_3).$$

Then

$$0 = a - a + a = a.$$

Therefore $a = 0$, and so μ^* vanishes on the corresponding 5 atoms as shown in the information diagram.



Structure of μ^* for $X_1 \to X_2 \to X_3 \to X_4$

1. The Markov subchain $X_1 \to X_2 \to X_3$ implies

$$0 = I(X_1; X_3|X_2) = I(X_1; X_3; X_4|X_2) + I(X_1; X_3|X_2, X_4).$$

Let $I(X_1; X_3|X_2, X_4) = a \geq 0$. Then

$$I(X_1; X_3; X_4|X_2) = -a.$$

2. The Markov subchain $X_1 \to X_2 \to X_4$ implies

$$0 = I(X_1; X_4|X_2) = I(X_1; X_3; X_4|X_2) + I(X_1; X_4|X_2, X_3).$$

Since $I(X_1; X_3; X_4|X_2) = -a$,

$$I(X_1; X_4|X_2, X_3) = a.$$

3. The Markov subchain $X_1 \to X_3 \to X_4$ implies

$$0 = I(X_1; X_4|X_3) = I(X_1; X_2; X_4|X_3) + I(X_1; X_4|X_2, X_3).$$

Since $I(X_1; X_4|X_2, X_3) = a$,

$$I(X_1; X_2; X_4|X_3) = -a.$$

4. The Markov subchain $X_2 \to X_3 \to X_4$ implies

$$0 = I(X_2; X_4|X_3) = I(X_1; X_2; X_4|X_3) + I(X_2; X_4|X_1, X_3).$$

Since $I(X_1; X_2; X_4|X_3) = -a$,

$$I(X_2; X_4|X_1, X_3) = a.$$

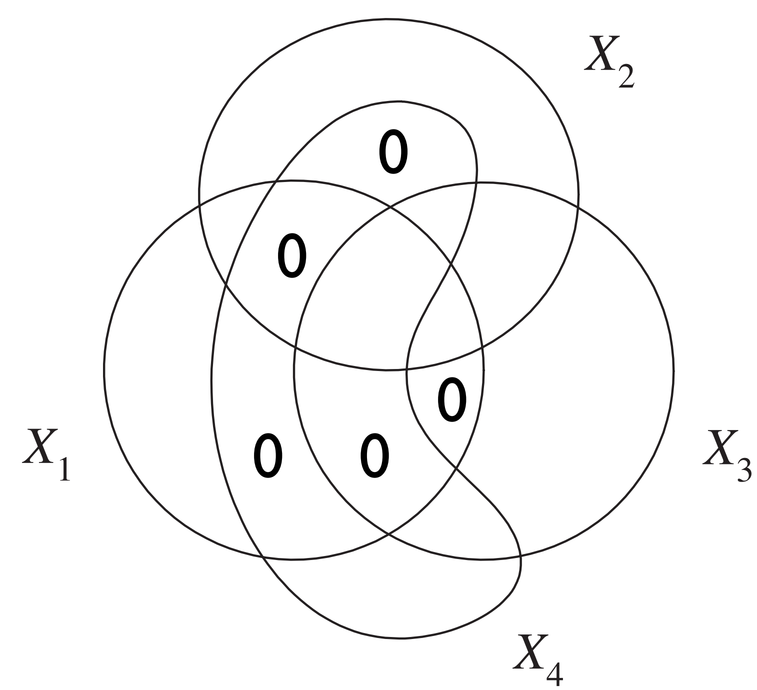
5. The Markov subchain $(X_1, X_2) \to X_3 \to X_4$ implies

$$0 = I(X_1, X_2; X_4|X_3) = I(X_1; X_4|X_2, X_3) + I(X_1; X_2; X_4|X_3) + I(X_2; X_4|X_1, X_3).$$

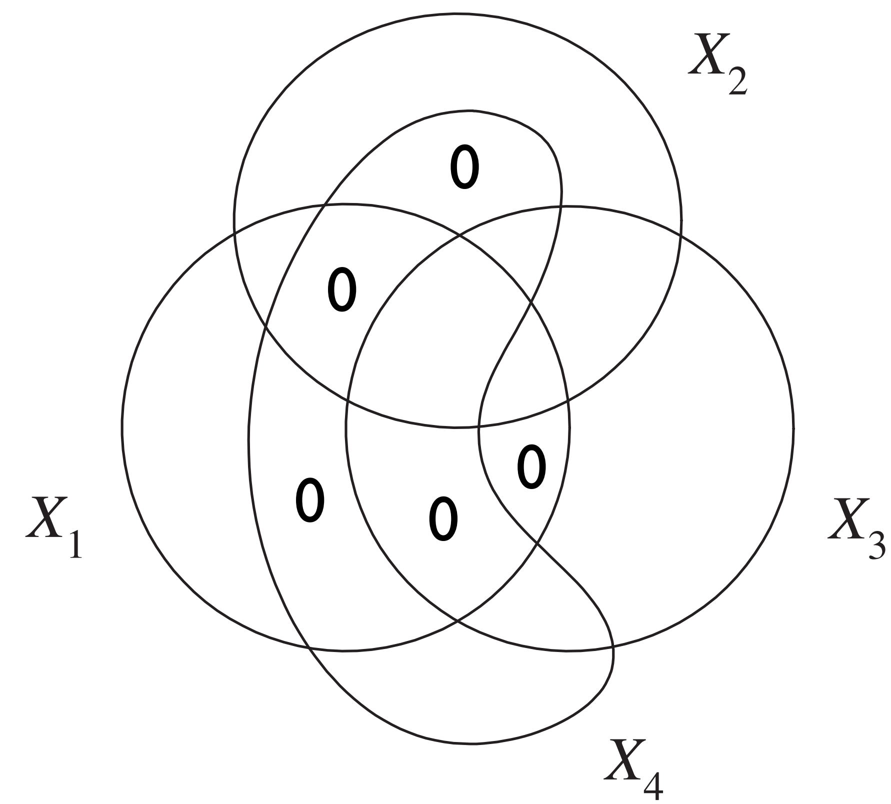
Then

$$0 = a - a + a = a.$$

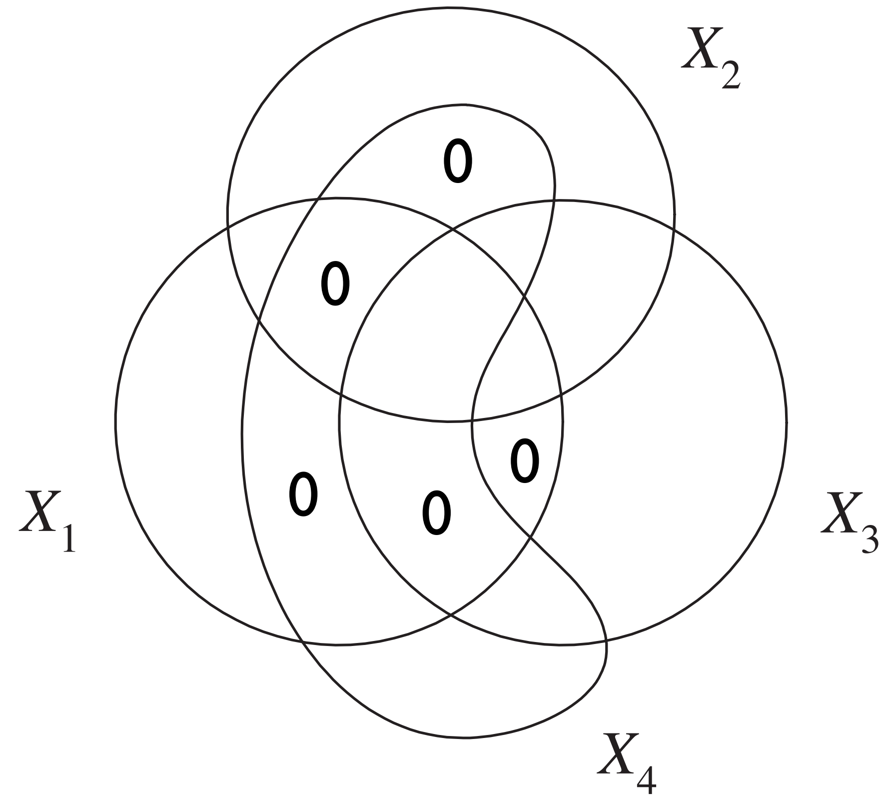
Therefore $a = 0$, and so μ^* vanishes on the corresponding 5 atoms as shown in the information diagram.



Nonnegatvity of μ^* for $X_1 \to X_2 \to X_3 \to X_4$



Nonnegatvity of μ^* for $X_1 \to X_2 \to X_3 \to X_4$

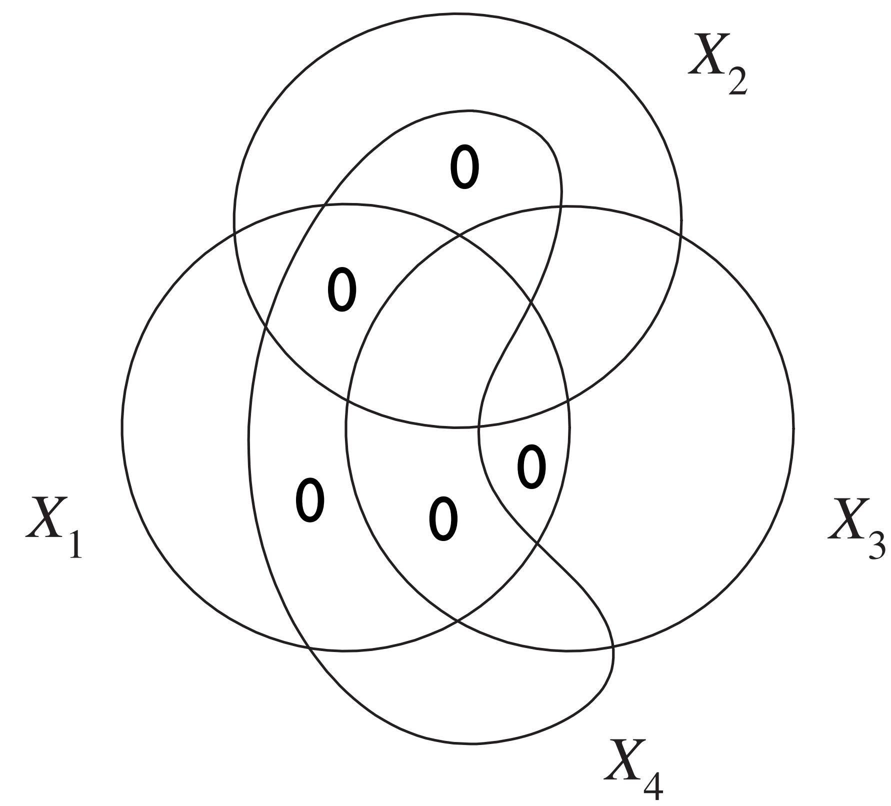


1. We have proved that μ^* vanishes on the 5 atoms shown in the information diagram.

Nonnegatvity of μ^* for $X_1 \to X_2 \to X_3 \to X_4$

1. We have proved that μ^* vanishes on the 5 atoms shown in the information diagram.

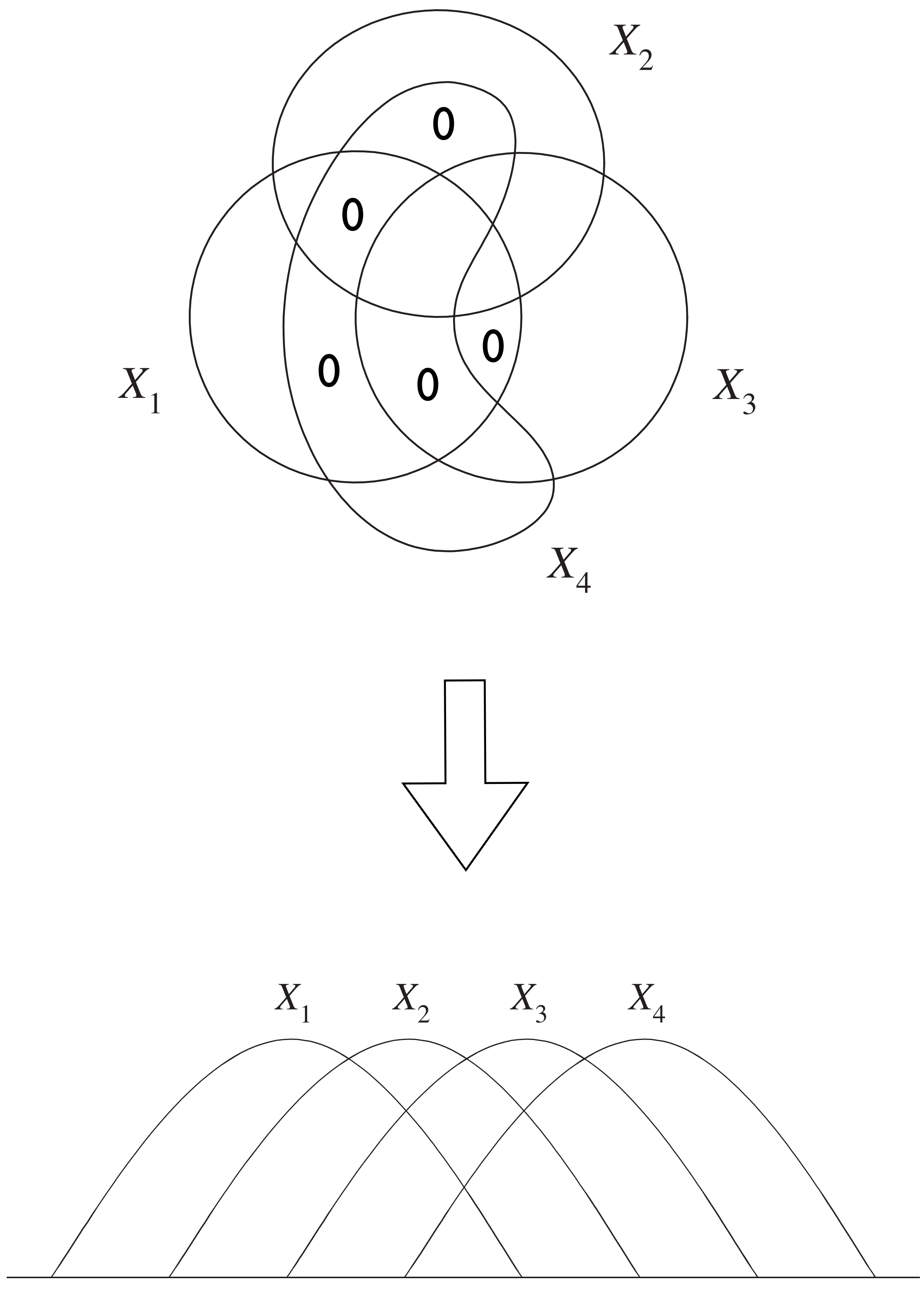
2. Suppress these atoms by setting them to $\emptyset$ to obtain the information diagram below.



Nonnegatvity of μ^* for $X_1 \to X_2 \to X_3 \to X_4$

1. We have proved that μ^* vanishes on the 5 atoms shown in the information diagram.

2. Suppress these atoms by setting them to $\emptyset$ to obtain the information diagram below.

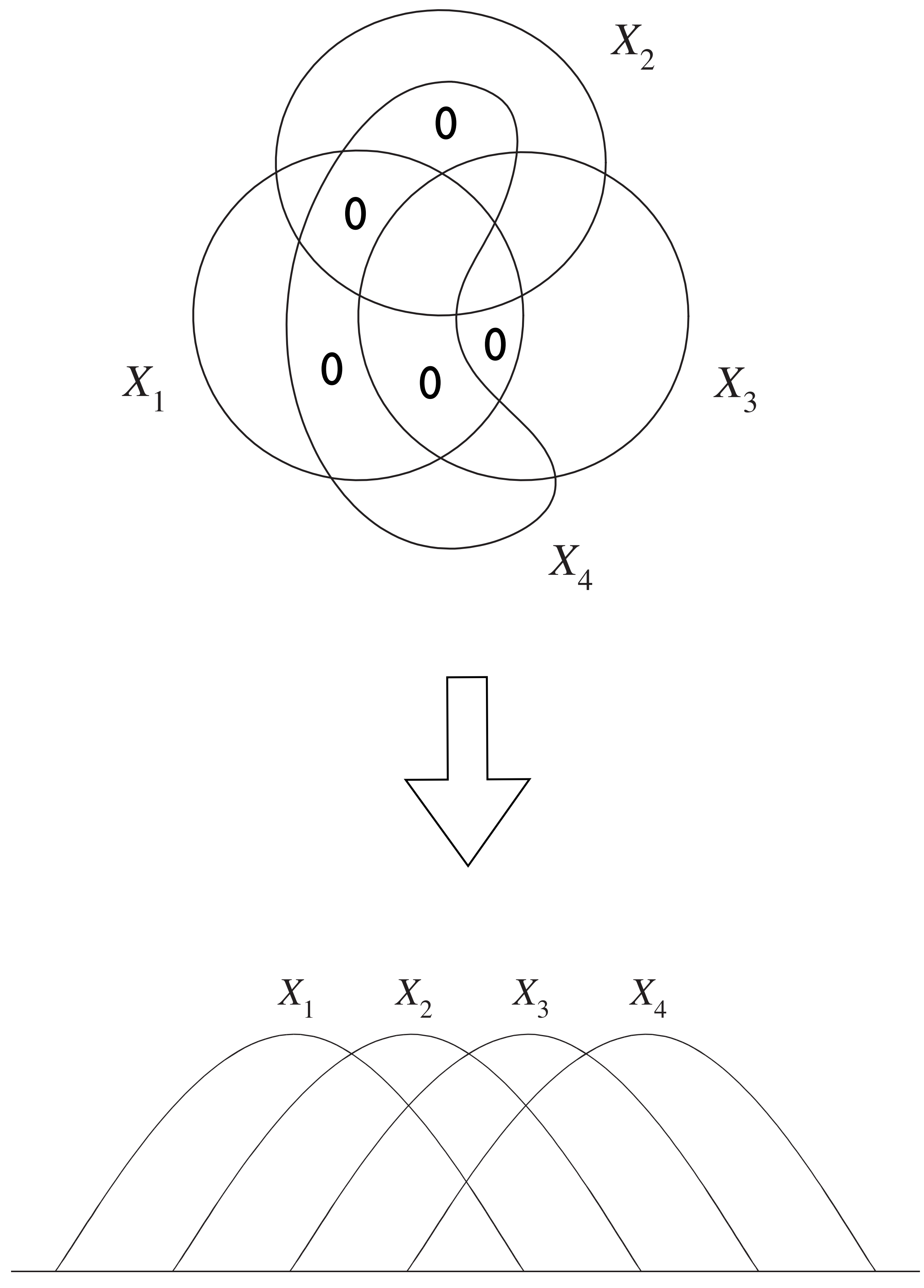


Nonnegatvity of μ^* for $X_1 \to X_2 \to X_3 \to X_4$

1. We have proved that μ^* vanishes on the 5 atoms shown in the information diagram.

2. Suppress these atoms by setting them to $\emptyset$ to obtain the information diagram below.

3. From this information diagram, it can readily be checked that the values of μ^* on the remaining



Nonnegatvity of μ^* for $X_1 \to X_2 \to X_3 \to X_4$

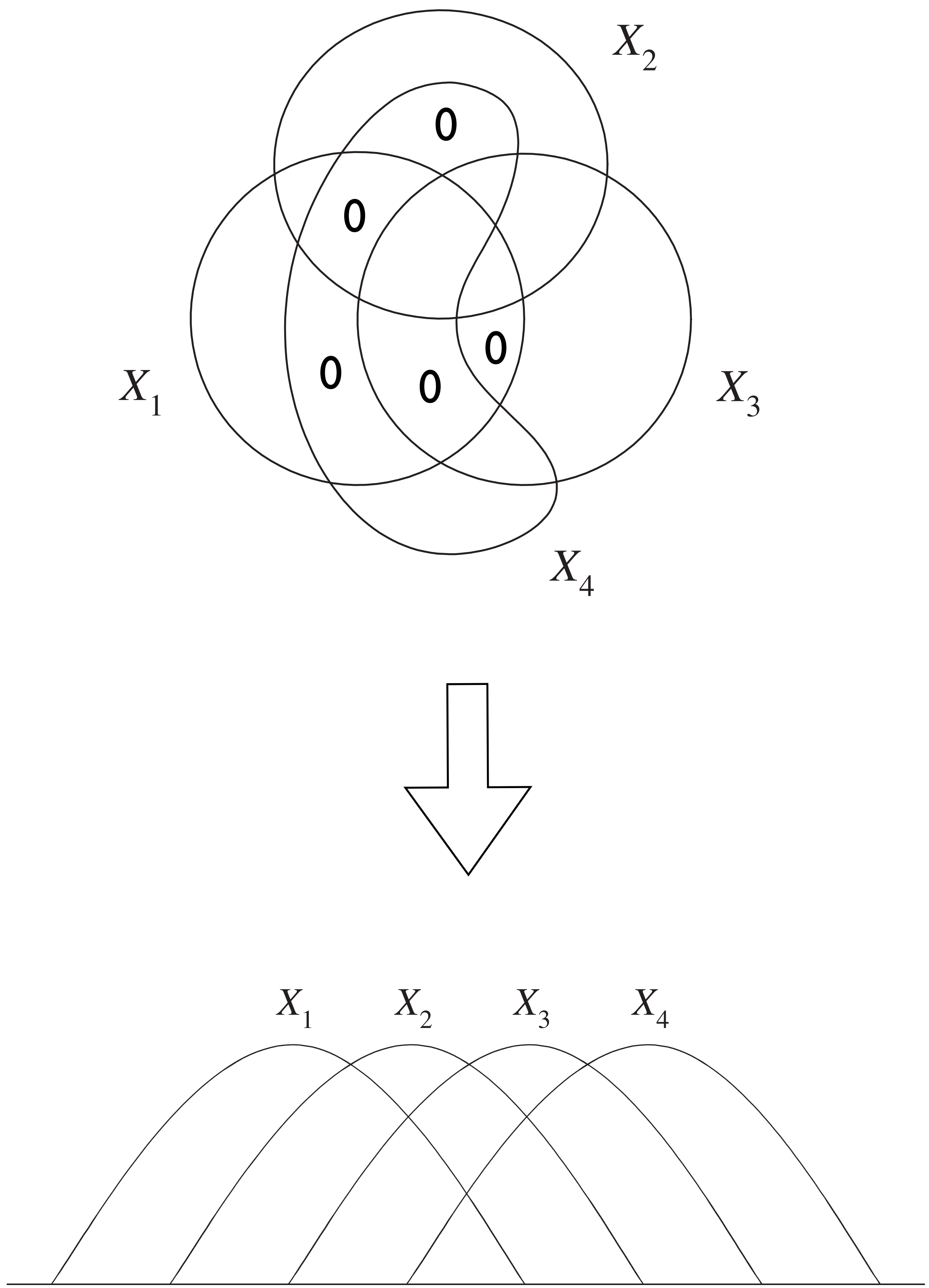
1. We have proved that μ^* vanishes on the 5 atoms shown in the information diagram.

2. Suppress these atoms by setting them to $\emptyset$ to obtain the information diagram below.

3. From this information diagram, it can readily be checked that the values of μ^* on the remaining

$$(2^4 - 1) - 5 = 10$$

nonempty atoms are equal to



Nonnegatvity of μ^* for $X_1 \to X_2 \to X_3 \to X_4$

1. We have proved that μ^* vanishes on the 5 atoms shown in the information diagram.

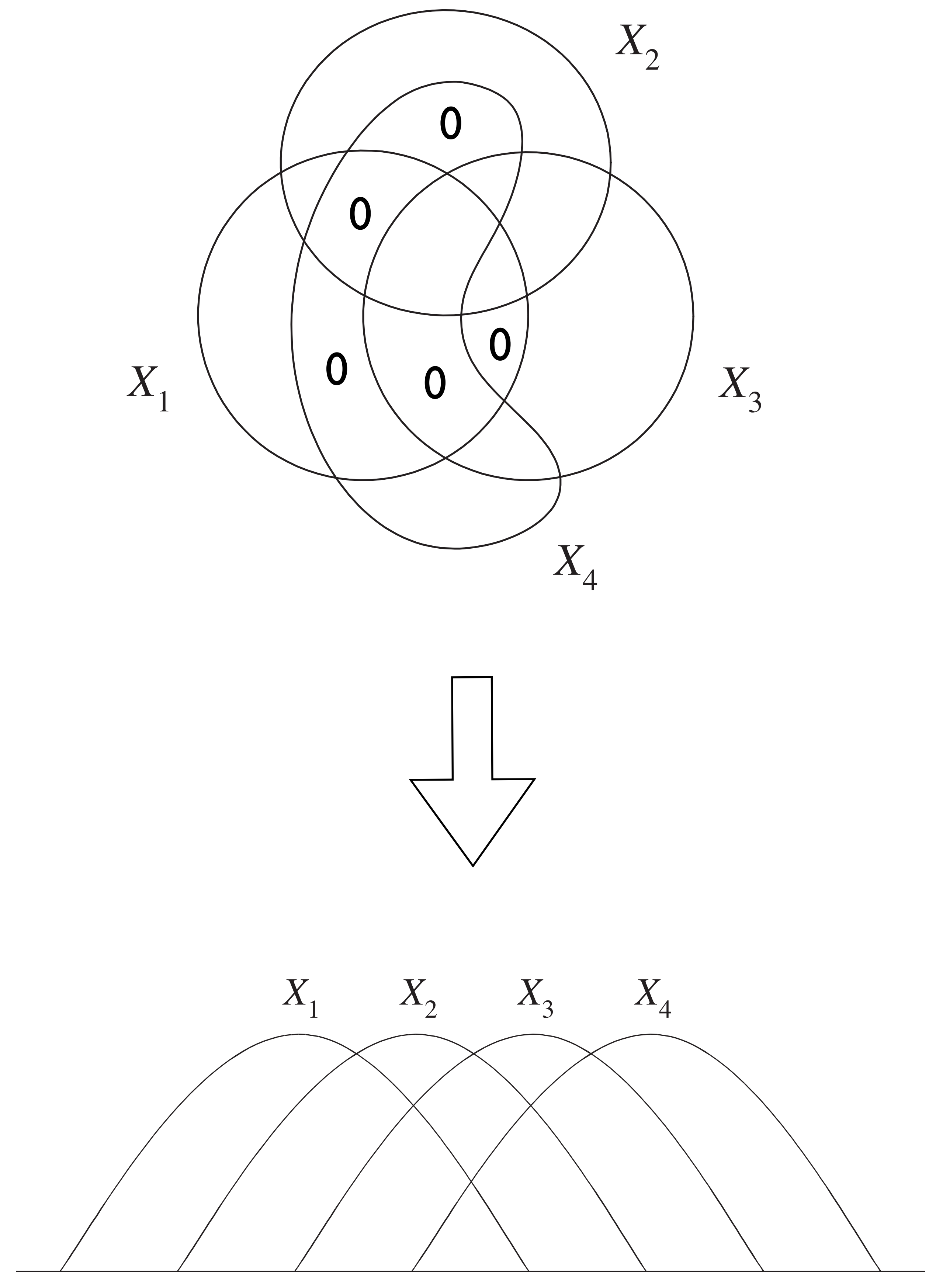
2. Suppress these atoms by setting them to $\emptyset$ to obtain the information diagram below.

3. From this information diagram, it can readily be checked that the values of μ^* on the remaining

$$(2^4 - 1) - 5 = 10$$

nonempty atoms are equal to

$$H(X_1|X_2, X_3, X_4),\ I(X_1; X_2|X_3, X_4)$$
$$I(X_1; X_3|X_4),\ I(X_1; X_4)$$
$$H(X_2|X_1, X_3, X_4),\ I(X_2; X_3|X_1; X_4)$$
$$I(X_2; X_4|X_1),\ H(X_3|X_1, X_2, X_4)$$
$$I(X_3; X_4|X_1, X_2),\ H(X_4|X_1, X_2, X_3).$$



Nonnegatvity of μ^* for $X_1 \to X_2 \to X_3 \to X_4$

1. We have proved that μ^* vanishes on the 5 atoms shown in the information diagram.

2. Suppress these atoms by setting them to $\emptyset$ to obtain the information diagram below.

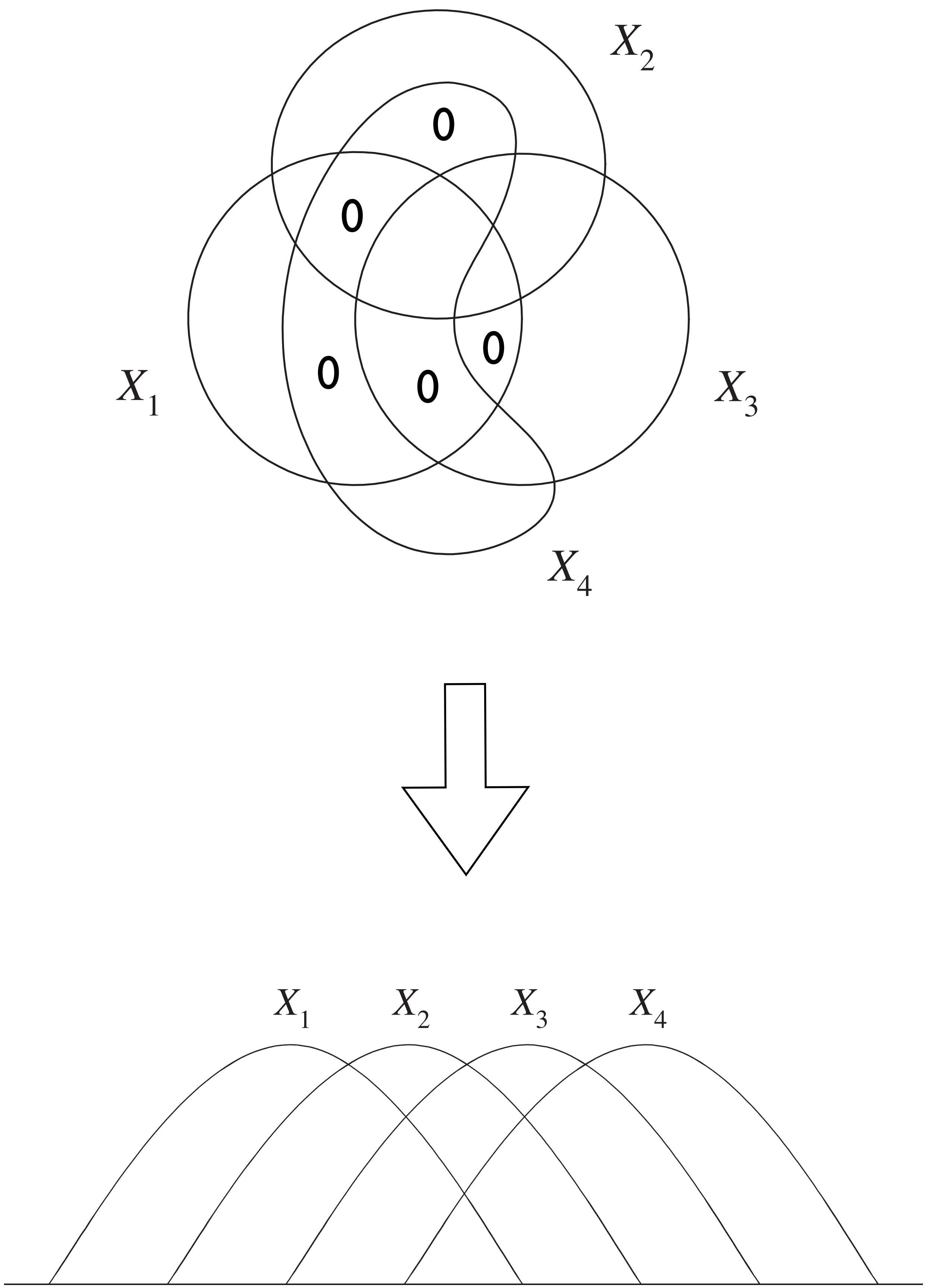
3. From this information diagram, it can readily be checked that the values of μ^* on the remaining

$$(2^4 - 1) - 5 = 10$$

nonempty atoms are equal to

$$H(X_1|X_2, X_3, X_4),\ I(X_1; X_2|X_3, X_4)$$
$$I(X_1; X_3|X_4),\ I(X_1; X_4)$$
$$H(X_2|X_1, X_3, X_4),\ I(X_2; X_3|X_1; X_4)$$
$$I(X_2; X_4|X_1),\ H(X_3|X_1, X_2, X_4)$$
$$I(X_3; X_4|X_1, X_2),\ H(X_4|X_1, X_2, X_3).$$

3. There are all Shannon's information measures which are always nonnegative. Therefore, μ^* is a measure.



Nonnegatvity of μ^* for $X_1 \to X_2 \to X_3 \to X_4$

1. We have proved that μ^* vanishes on the 5 atoms shown in the information diagram.

2. Suppress these atoms by setting them to $\emptyset$ to obtain the information diagram below.

3. From this information diagram, it can readily be checked that the values of μ^* on the remaining

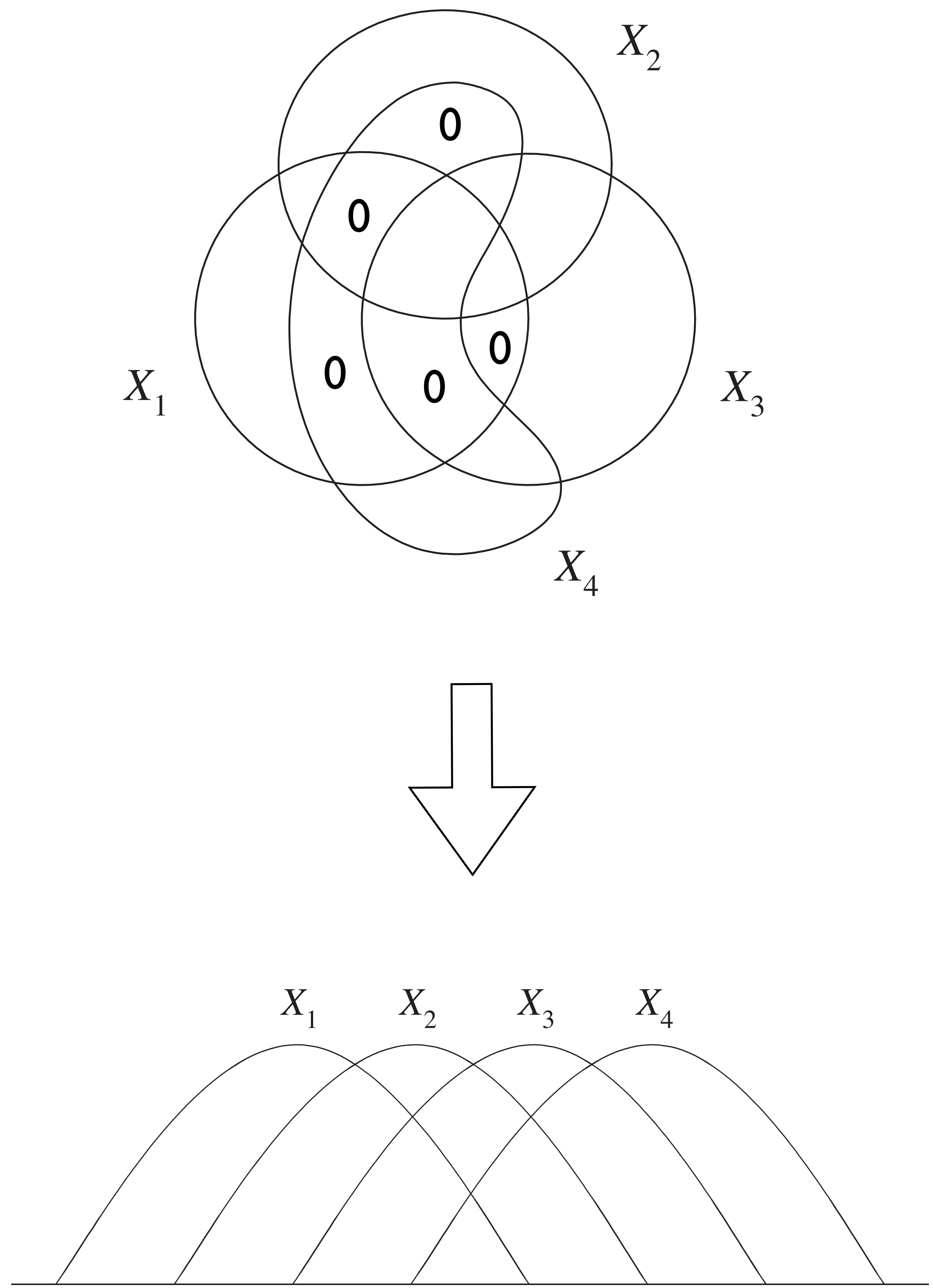
$$(2^4 - 1) - 5 = 10$$

nonempty atoms are equal to

$$H(X_1|X_2, X_3, X_4),\ I(X_1; X_2|X_3, X_4)$$
$$I(X_1; X_3|X_4),\ I(X_1; X_4)$$
$$H(X_2|X_1, X_3, X_4),\ I(X_2; X_3|X_1; X_4)$$
$$I(X_2; X_4|X_1),\ H(X_3|X_1, X_2, X_4)$$
$$I(X_3; X_4|X_1, X_2),\ H(X_4|X_1, X_2, X_3).$$

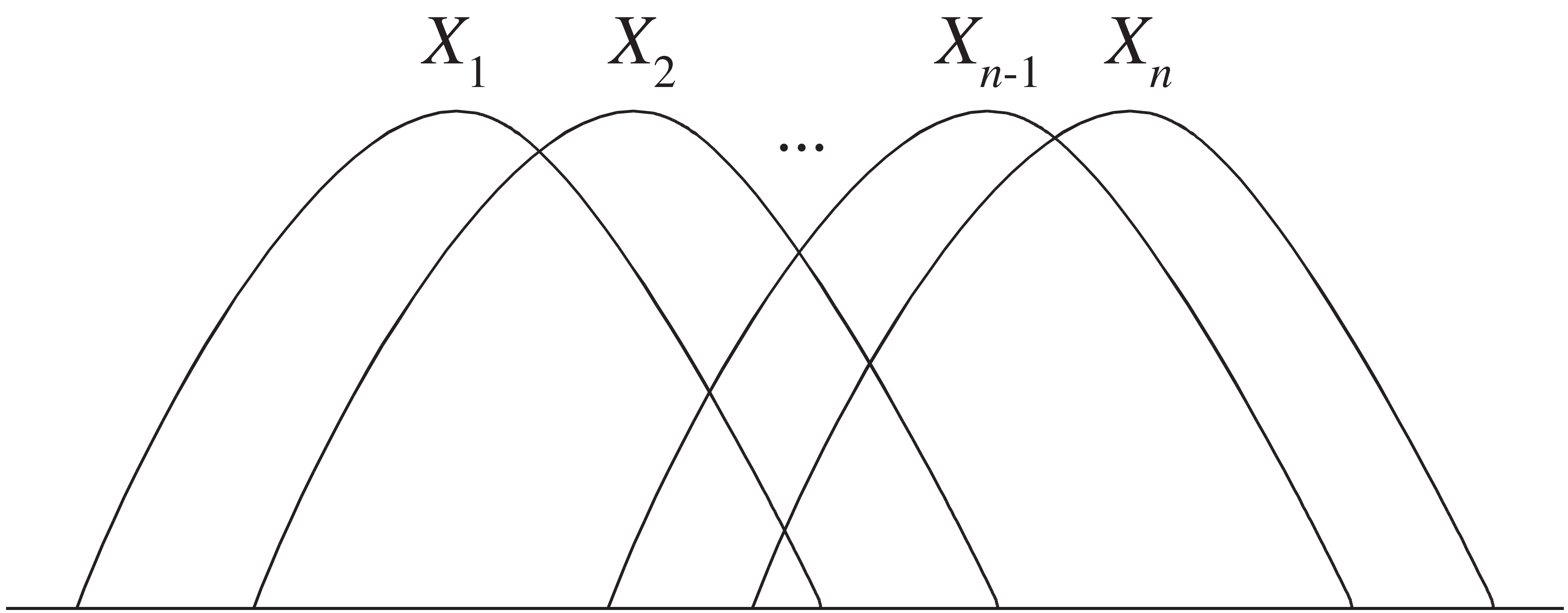
3. There are all Shannon's information measures which are always nonnegative. Therefore, μ^* is a measure.

Exercise: Identify these 10 atoms in the information diagram at the bottom.



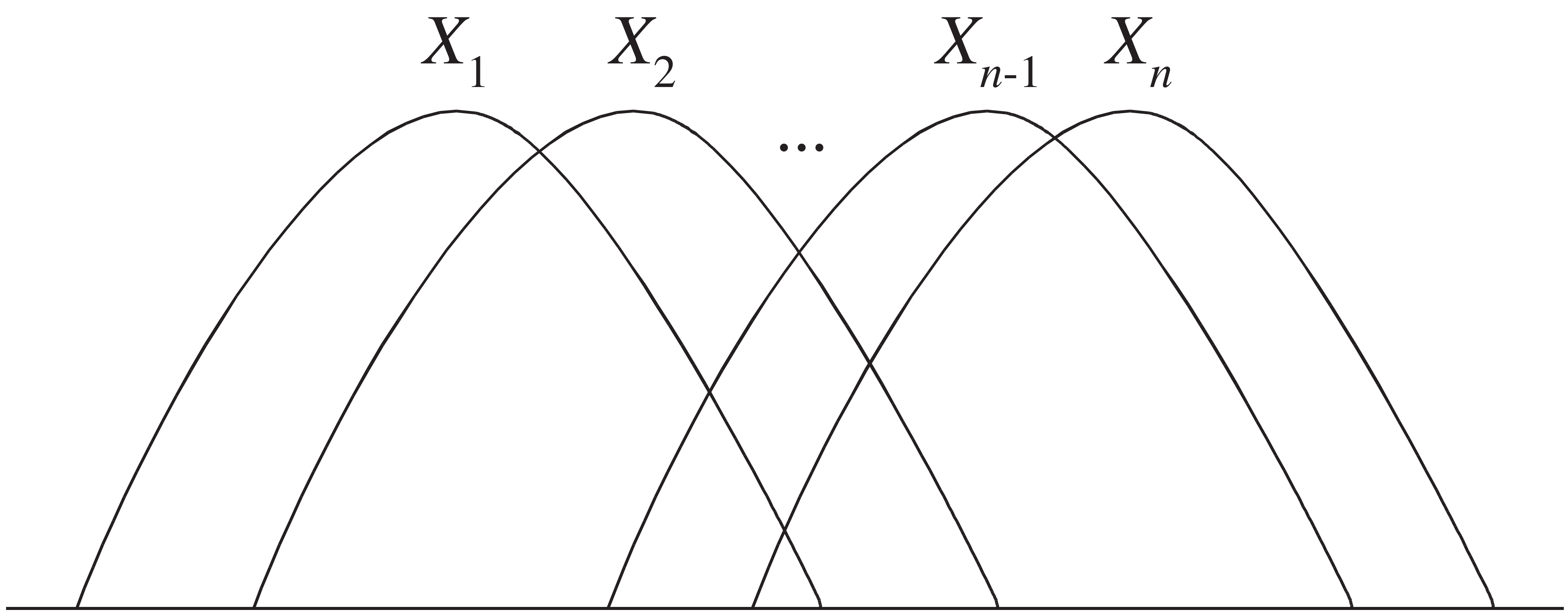
- For a general n, the information diagram can be displayed in two dimensions because certain atoms can be suppressed.

- For a general n, the information diagram can be displayed in two dimensions because certain atoms can be suppressed.



- For a general n, the information diagram can be displayed in two dimensions because certain atoms can be suppressed.

- The values of μ^* on the remaining atoms correspond to Shannon's information measures and hence are nonnegative. Thus, μ^* is a measure.



- For a general n, the information diagram can be displayed in two dimensions because certain atoms can be suppressed.
- The values of μ^* on the remaining atoms correspond to Shannon's information measures and hence are nonnegative. Thus, μ^* is a measure.
- See Ch. 12 for a detailed discussion in the context of Markov random field.