

Entropia di sistemi di informazione incompleti

Gianpiero Cattaneo, Davide Ciucci, Daniela Bianucci

Dipartimento di Informatica, Sistemistica e Comunicazione
Università di Milano Bicocca

17 luglio 2006/ Varese

IS: example

Flat	Price	Rooms	Down-Town	Furniture
f_1	high	2	yes	no
f_2	high	1	yes	no
f_3	high	2	yes	no
f_4	low	1	no	no
f_5	low	1	no	no
f_6	medium	1	yes	yes

Definition

An **Information System** (IS) is a structure $\mathcal{K}(X) = \langle X, Att, val, F \rangle$ where

- the **universe** X is a non empty set of **objects**

Definition

An **Information System** (IS) is a structure $\mathcal{K}(X) = \langle X, Att, val, F \rangle$ where

- the **universe** X is a non empty set of **objects**
- Att is a non empty set of **attributes**

Definition

An **Information System** (IS) is a structure $\mathcal{K}(X) = \langle X, Att, val, F \rangle$ where

- the **universe** X is a non empty set of **objects**
- Att is a non empty set of **attributes**
- val is the set of all **possible values** that can be observed for an attribute a from Att

Definition

An **Information System** (IS) is a structure $\mathcal{K}(X) = \langle X, Att, val, F \rangle$ where

- the **universe** X is a non empty set of **objects**
- Att is a non empty set of **attributes**
- val is the set of all **possible values** that can be observed for an attribute a from Att
- F (called the **information map**) is a function

$$F : X \times Att \rightarrow val$$

which associates to any pair $x \in X$ and $a \in Att$
the value $F(x, a) \in val$ assumed by the attribute a for the object x

Definition (Indiscernibilità)

Dato un insieme di attributi $A \in Att$

Definition (Indiscernibilità)

Dato un insieme di attributi $A \in Att$
due oggetti $x, y \in U$ sono detti **indiscernibili** rispetto a A se

Definition (Indiscernibilità)

Dato un insieme di attributi $A \in Att$
due oggetti $x, y \in U$ sono detti **indiscernibili** rispetto a A se

$$\forall a \in A \quad F(a, x) = F(a, y)$$

In questo caso scriviamo $xI_A y$

Definition (Indiscernibilità)

Dato un insieme di attributi $A \in Att$
due oggetti $x, y \in U$ sono detti **indiscernibili** rispetto a A se

$$\forall a \in A \quad F(a, x) = F(a, y)$$

In questo caso scriviamo $xI_A y$

I_A è una relazione di equivalenza che partiziona U in classi di equivalenza

$$C_A(x) := \{y \in U : xI_A y\}$$

Classi Equivalenza: esempio

Classi Equivalenza: esempio

Paziente	Pressione	Mal di Testa	Temperatura	Dolori Muscolari
P1	Normale	si	38–39	si
P2	Alta	no	36–37	si
P3	Alta	no	36–37	si
P4	Bassa	si	35–36	no
P5	Normale	si	36–37	si

Classi Equivalenza: esempio

Paziente	Pressione	Mal di Testa	Temperatura	Dolori Muscolari
P1	Normale	si	38–39	si
P2	Alta	no	36–37	si
P3	Alta	no	36–37	si
P4	Bassa	si	35–36	no
P5	Normale	si	36–37	si

Classi Equivalenza: esempio

Paziente	Pressione	Mal di Testa	Temperatura	Dolori Muscolari
P1	Normale	si	38–39	si
P2	Alta	no	36–37	si
P3	Alta	no	36–37	si
P4	Bassa	si	35–36	no
P5	Normale	si	36–37	si

Classi Equivalenza: esempio

Paziente	Pressione	Mal di Testa	Temperatura	Dolori Muscolari
P1	Normale	si	38–39	si
P2	Alta	no	36–37	si
P3	Alta	no	36–37	si
P4	Bassa	si	35–36	no
P5	Normale	si	36–37	si

Le **classi di equivalenza** C_i sono {P1}, {P2,P3}, {P4}, {P5}

Entropia di una partizione

A insieme di attributi

Entropia di una partizione

A insieme di attributi

C_i classe di equivalenza: rappresenta l'“evento” di osservare una determinata tupla di valori

Entropia di una partizione

A insieme di attributi

C_i classe di equivalenza: rappresenta l'“evento” di osservare una determinata tupla di valori

Probabilità dell'evento C_i : $p(C_i) = \frac{|C_i|}{|X|}$

Entropia di una partizione

A insieme di attributi

C_i classe di equivalenza: rappresenta l'“evento” di osservare una determinata tupla di valori

Probabilità dell'evento C_i : $p(C_i) = \frac{|C_i|}{|X|}$

Definition

Entropia della partizione generata da A

$$H(A) = - \sum_{i=1}^n \frac{|C_i|}{|X|} \log_2 \frac{|C_i|}{|X|}$$

Entropia di una partizione

A insieme di attributi

C_i classe di equivalenza: rappresenta l'“evento” di osservare una determinata tupla di valori

Probabilità dell'evento C_i : $p(C_i) = \frac{|C_i|}{|X|}$

Definition

Entropia della partizione generata da A

$$H(A) = - \sum_{i=1}^n \frac{|C_i|}{|X|} \log_2 \frac{|C_i|}{|X|} = \log_2 |X| - \frac{1}{|X|} \sum_{i=1}^n |C_i| \log_2 |C_i|$$

Entropia di una partizione

A insieme di attributi

C_i classe di equivalenza: rappresenta l'“evento” di osservare una determinata tupla di valori

Probabilità dell'evento C_i : $p(C_i) = \frac{|C_i|}{|X|}$

Definition

Entropia della partizione generata da A

$$H(A) = - \sum_{i=1}^n \frac{|C_i|}{|X|} \log_2 \frac{|C_i|}{|X|} = \log_2 |X| - \frac{1}{|X|} \sum_{i=1}^n |C_i| \log_2 |C_i|$$

co-entropia $E(A) = \frac{1}{|X|} \sum_{i=1}^n |C_i| \log_2 |C_i|$

Entropia di una partizione

A insieme di attributi

C_i classe di equivalenza: rappresenta l'“evento” di osservare una determinata tupla di valori

Probabilità dell'evento C_i : $p(C_i) = \frac{|C_i|}{|X|}$

Definition

Entropia della partizione generata da A

$$H(A) = - \sum_{i=1}^n \frac{|C_i|}{|X|} \log_2 \frac{|C_i|}{|X|} = \log_2 |X| - \frac{1}{|X|} \sum_{i=1}^n |C_i| \log_2 |C_i|$$

co-entropia $E(A) = \frac{1}{|X|} \sum_{i=1}^n |C_i| \log_2 |C_i|$

$$H(A) + E(A) = \log_2(|X|)$$

Partizione **discreta**: $\pi_d = \{\{x_1\}, \{x_2\}, \dots, \{x_n\}\}$

Partizione **discreta**: $\pi_d = \{\{x_1\}, \{x_2\}, \dots, \{x_n\}\}$

$$H(\pi_d) = \log |X| \quad E(\pi_d) = 0$$

Casi particolari

Partizione **discreta**: $\pi_d = \{\{x_1\}, \{x_2\}, \dots, \{x_n\}\}$

$$H(\pi_d) = \log |X| \quad E(\pi_d) = 0$$

Partizione **banale**: $\pi_t = \{X\}$

Casi particolari

Partizione **discreta**: $\pi_d = \{\{x_1\}, \{x_2\}, \dots, \{x_n\}\}$

$$H(\pi_d) = \log |X| \quad E(\pi_d) = 0$$

Partizione **banale**: $\pi_t = \{X\}$

$$H(\pi_t) = 0 \quad E(\pi_d) = \log |X|$$

Casi particolari

Partizione **discreta**: $\pi_d = \{\{x_1\}, \{x_2\}, \dots, \{x_n\}\}$

$$H(\pi_d) = \log |X| \quad E(\pi_d) = 0$$

Partizione **banale**: $\pi_t = \{X\}$

$$H(\pi_t) = 0 \quad E(\pi_d) = \log |X|$$

Data una generica partizione π

Casi particolari

Partizione **discreta**: $\pi_d = \{\{x_1\}, \{x_2\}, \dots, \{x_n\}\}$

$$H(\pi_d) = \log |X| \quad E(\pi_d) = 0$$

Partizione **banale**: $\pi_t = \{X\}$

$$H(\pi_t) = 0 \quad E(\pi_d) = \log |X|$$

Data una generica partizione π

$$0 = H(\pi_t) \leq H(\pi) \leq H(\pi_d) = \log |X|$$

Ordine tra partizioni

Sull'insieme di tutte le partizioni $\Pi(X)$ di un universo X è possibile definire una **relazione d'ordine**

$$\pi_1 \preceq \pi_2 \quad \text{iff} \quad \forall C_i \in \pi_1, \exists D_j \in \pi_2 : C_i \subseteq D_j$$

Ordine tra partizioni

Sull'insieme di tutte le partizioni $\Pi(X)$ di un universo X è possibile definire una **relazione d'ordine**

$$\pi_1 \preceq \pi_2 \quad \text{iff} \quad \forall C_i \in \pi_1, \exists D_j \in \pi_2 : C_i \subseteq D_j$$

$$\pi_1 \ll \pi_2 \quad \text{iff} \quad \forall D_j \in \pi_2, \exists \{C_{i_1}, C_{i_2}, \dots, C_{i_p}\} \subseteq \pi_1 : D_j = C_{i_1} \cup C_{i_2} \cup \dots \cup C_{i_p}$$

Ordine tra partizioni

Sull'insieme di tutte le partizioni $\Pi(X)$ di un universo X è possibile definire una **relazione d'ordine**

$$\pi_1 \preceq \pi_2 \quad \text{iff} \quad \forall C_i \in \pi_1, \exists D_j \in \pi_2 : C_i \subseteq D_j$$

$$\pi_1 \ll \pi_2 \quad \text{iff} \quad \forall D_j \in \pi_2, \exists \{C_{i_1}, C_{i_2}, \dots, C_{i_p}\} \subseteq \pi_1 : D_j = C_{i_1} \cup C_{i_2} \cup \dots \cup C_{i_p}$$

$$\forall \pi \in \Pi(X), \quad \pi_d \prec \pi \prec \pi_t$$

Ordine tra partizioni

Sull'insieme di tutte le partizioni $\Pi(X)$ di un universo X è possibile definire una **relazione d'ordine**

$$\pi_1 \preceq \pi_2 \quad \text{iff} \quad \forall C_i \in \pi_1, \exists D_j \in \pi_2 : C_i \subseteq D_j$$

$$\pi_1 \ll \pi_2 \quad \text{iff} \quad \forall D_j \in \pi_2, \exists \{C_{i_1}, C_{i_2}, \dots, C_{i_p}\} \subseteq \pi_1 : D_j = C_{i_1} \cup C_{i_2} \cup \dots \cup C_{i_p}$$

$$\forall \pi \in \Pi(X), \quad \pi_d \prec \pi \prec \pi_t$$

$$0 = H(\pi_d) \leq H(\pi) \leq H(\pi_t) = \log |X|$$

Theorem

$\pi_1 \preceq \pi_2$ *implica* $H(\pi_2) \leq H(\pi_1)$

Theorem

$\pi_1 \preceq \pi_2$ *implica* $H(\pi_2) \leq H(\pi_1)$

$\pi_1 \preceq \pi_2$ *implica* $E(\pi_1) \leq E(\pi_2)$

Theorem

$\pi_1 \preceq \pi_2$ *implica* $H(\pi_2) \leq H(\pi_1)$

$\pi_1 \preceq \pi_2$ *implica* $E(\pi_1) \leq E(\pi_2)$

Tornando ai sistemi di informazione...

Theorem

$A \subseteq B \subseteq \text{Att}$ *implica* $\pi(B) \preceq \pi(A)$

Theorem

$\pi_1 \preceq \pi_2$ *implica* $H(\pi_2) \leq H(\pi_1)$

$\pi_1 \preceq \pi_2$ *implica* $E(\pi_1) \leq E(\pi_2)$

Tornando ai sistemi di informazione...

Theorem

$A \subseteq B \subseteq \text{Att}$ *implica* $\pi(B) \preceq \pi(A)$

Quindi

$$H(\pi(B)) \leq H(\pi(A))$$

$$E(\pi(A)) \leq E(\pi(B))$$

Sistemi Incompleti: Esempio

Sistemi Incompleti: Esempio

Flat	Price	Rooms	Down-Town	Furniture
1	high	2	yes	*
2	high	*	yes	no
3	*	2	yes	no
4	low	*	no	no
5	low	1	*	no
6	*	*	yes	*
7	h	*	*	no
8	l	1	*	*
9	h	2	yes	no
10	l	1	no	*
11	*	2	yes	*

Definition

An **Incomplete Information System** (IIS) is a structure

$\mathcal{K}(X) = \langle X, Att, val(X), F \rangle$ where

$F : X \times Att \mapsto val(X)$ is **partially** defined on a subset $D(\mathcal{K})$ of $X \times Att$

Definition

An **Incomplete Information System** (IIS) is a structure

$\mathcal{K}(X) = \langle X, Att, val(X), F \rangle$ where

$F : X \times Att \mapsto val(X)$ is **partially** defined on a subset $D(\mathcal{K})$ of $X \times Att$

$\forall a \in Att, \quad X_a$ insieme degli oggetti definiti su a

$$X_a := \{x \in X : (x, a) \in D(\mathcal{K})\}$$

Definition

An **Incomplete Information System** (IIS) is a structure

$\mathcal{K}(X) = \langle X, Att, val(X), F \rangle$ where

$F : X \times Att \mapsto val(X)$ is **partially** defined on a subset $D(\mathcal{K})$ of $X \times Att$

$\forall a \in Att, \quad X_a$ insieme degli oggetti definiti su a

$$X_a := \{x \in X : (x, a) \in D(\mathcal{K})\}$$

Ipotesi: $\cup_a X_a = X$

Come definire entropia? Caso 1

Come definire entropia? Caso 1

In genere si definisce la **relazione di similarità**

Come definire entropia? Caso 1

In genere si definisce la **relazione di similarità**

Let $x, y \in X$, then $xS_A y$ iff $\forall a_i \in A$,
either $f_{a_i}(x) = f_{a_i}(y)$ or $f_{a_i}(x) = *$ or $f_{a_i}(y) = *$

Come definire entropia? Caso 1

In genere si definisce la **relazione di similarità**

Let $x, y \in X$, then $xS_A y$ iff $\forall a_i \in A$,
either $f_{a_i}(x) = f_{a_i}(y)$ or $f_{a_i}(x) = *$ or $f_{a_i}(y) = *$

Classi di similarità $S_A(x) := \{y \in X : xS_A y\}$

Come definire entropia? Caso 1

In genere si definisce la **relazione di similarità**

Let $x, y \in X$, then $xS_A y$ iff $\forall a_i \in A$,
either $f_{a_i}(x) = f_{a_i}(y)$ or $f_{a_i}(x) = *$ or $f_{a_i}(y) = *$

Classi di similarità $S_A(x) := \{y \in X : xS_A y\}$

L'unione delle classi di similarità è un **covering** (ricoprimento)
dell'universo X

$$\cup_x S_A(x) = X$$

in genere non una partizione

Sistemi Incompleti: Esempio

Sistemi Incompleti: Esempio

Flat	Price	Rooms	Down-Town	Furniture
1	high	2	yes	*
2	high	*	yes	no
3	*	2	yes	no
4	low	*	no	no
5	low	1	*	no
6	medium	*	yes	*

Classi di similarità: $\{1, 2, 3\}$, $\{1, 2, 3, 6\}$, $\{4, 5\}$, $\{3, 6\}$

Come definire entropia? Caso 2

Come definire entropia? Caso 2

Siano a, b due attributi

$val^*(a)$, $val^*(b)$ i valori che possono assumere a, b più il valore $*$

Come definire entropia? Caso 2

Siano a, b due attributi

$val^*(a)$, $val^*(b)$ i valori che possono assumere a, b più il valore $*$

Definiamo $f_{a,b} : X_a \cup X_b \mapsto val^*(a) \times val^*(b)$ come

Come definire entropia? Caso 2

Siano a, b due attributi

$val^*(a)$, $val^*(b)$ i valori che possono assumere a, b più il valore $*$

Definiamo $f_{a,b} : X_a \cup X_b \mapsto val^*(a) \times val^*(b)$ come

$$f_{a,b}(x) := \begin{cases} (f_a(x), f_b(x)) & x \in X_a \cap X_b \\ (f_a(x), *) & x \in X_a \cap (X_b)^c \\ (*, f_b(x)) & x \in (X_a)^c \cap X_b \end{cases}$$

Come definire entropia? Caso 2

Siano a, b due attributi

$val^*(a)$, $val^*(b)$ i valori che possono assumere a, b più il valore $*$

Definiamo $f_{a,b} : X_a \cup X_b \mapsto val^*(a) \times val^*(b)$ come

$$f_{a,b}(x) := \begin{cases} (f_a(x), f_b(x)) & x \in X_a \cap X_b \\ (f_a(x), *) & x \in X_a \cap (X_b)^c \\ (*, f_b(x)) & x \in (X_a)^c \cap X_b \end{cases}$$

Generalizzando a un insieme di attributi A

$$f_A : X_A \mapsto val^*(A) \text{ dove } val^*(A) = \prod_{a_i \in A} val^*(a_i)$$

Come definire entropia? Caso 2

Siano a, b due attributi

$val^*(a)$, $val^*(b)$ i valori che possono assumere a, b più il valore $*$

Definiamo $f_{a,b} : X_a \cup X_b \mapsto val^*(a) \times val^*(b)$ come

$$f_{a,b}(x) := \begin{cases} (f_a(x), f_b(x)) & x \in X_a \cap X_b \\ (f_a(x), *) & x \in X_a \cap (X_b)^c \\ (*, f_b(x)) & x \in (X_a)^c \cap X_b \end{cases}$$

Generalizzando a un insieme di attributi A

$f_A : X_A \mapsto val^*(A)$ dove $val^*(A) = \prod_{a_i \in A} val^*(a_i)$

Il **granulo** $f_A^{-1}(\vec{\alpha}) = \{x \in X_A : f_A(x) = \vec{\alpha}\}$

Come definire entropia? Caso 2

Siano a, b due attributi

$val^*(a)$, $val^*(b)$ i valori che possono assumere a, b più il valore $*$

Definiamo $f_{a,b} : X_a \cup X_b \mapsto val^*(a) \times val^*(b)$ come

$$f_{a,b}(x) := \begin{cases} (f_a(x), f_b(x)) & x \in X_a \cap X_b \\ (f_a(x), *) & x \in X_a \cap (X_b)^c \\ (*, f_b(x)) & x \in (X_a)^c \cap X_b \end{cases}$$

Generalizzando a un insieme di attributi A

$$f_A : X_A \mapsto val^*(A) \text{ dove } val^*(A) = \prod_{a_i \in A} val^*(a_i)$$

$$\text{Il granulo } f_A^{-1}(\vec{\alpha}) = \{x \in X_A : f_A(x) = \vec{\alpha}\}$$

$\{f_A^{-1}(\vec{\alpha}) : \vec{\alpha} \in val^*(A)\}$ è una partizione di X_A

Caso 2

$$p(A_{\vec{\alpha}}) = \frac{|A_{\vec{\alpha}}|}{|X_{\mathcal{A}}|} \quad p(X^*) = 0$$

$$p(A_{\vec{\alpha}}) = \frac{|A_{\vec{\alpha}}|}{|X_{\mathcal{A}}|} \quad p(X^*) = 0$$

$$p(A_{\vec{\alpha}}) = \frac{|A_{\vec{\alpha}}|}{|X_{\mathcal{A}}|} \quad p(X^*) = 0$$

Definition (Entropia)

Entropia

$$H(\mathcal{A}) = - \sum_{\vec{\alpha}} p(A_{\vec{\alpha}}) \log p(A_{\vec{\alpha}}) = \log |X_{\mathcal{A}}| - \frac{1}{|X_{\mathcal{A}}|} \sum_{\vec{\alpha}} |A_{\alpha}| \log |A_{\alpha}|$$

Co-entropy

$$E(\mathcal{A}) = \frac{1}{|X_{\mathcal{A}}|} \sum_{\vec{\alpha} \in V_{\mathcal{A}}^*} |A_{\alpha}| \log |A_{\alpha}|$$

Caso 2 - ordine

Se $A \subseteq B$ allora $X_A \subseteq X_B$ e $\pi(B) \ll \pi(A)$
ma in generale **non vale** $H(A) \leq H(B)$

Caso 2 - ordine

Se $A \subseteq B$ allora $X_A \subseteq X_B$ e $\pi(B) \ll \pi(A)$
ma in generale **non vale** $H(A) \leq H(B)$

Problema: entropia dipende da $|X_{\mathcal{A}}|$

Caso 2 - ordine

Se $A \subseteq B$ allora $X_A \subseteq X_B$ e $\pi(B) \ll \pi(A)$
ma in generale **non vale** $H(A) \leq H(B)$

Problema: entropia dipende da $|X_{\mathcal{A}}|$

Soluzione

$$H^*(\mathcal{A}) = \log |X| - \frac{1}{|X|} \sum_{\vec{\alpha}} |A_{\alpha}| \log |A_{\alpha}|$$

$$E^*(\mathcal{A}) = \frac{1}{|X|} \sum_{\vec{\alpha} \in V_{\mathcal{A}}^*} |A_{\alpha}| \log |A_{\alpha}|$$

Idea: generalizzare le definizioni date nel caso delle partizioni

Caso 1 - soluzione 1

Idea: generalizzare le definizioni date nel caso delle partizioni

Partizione $\pi = \{A_1, \dots, A_n\}$	Covering $\gamma = \{B_1, \dots, B_n\}$
$\forall x \in X, \sum_{i=1}^n \chi_{A_i}(x) = 1$	

Caso 1 - soluzione 1

Idea: generalizzare le definizioni date nel caso delle partizioni

Partizione $\pi = \{A_1, \dots, A_n\}$	Covering $\gamma = \{B_1, \dots, B_n\}$
$\forall x \in X, \sum_{i=1}^n \chi_{A_i}(x) = 1$	$n(x) = \sum_{i=1}^n \chi_{B_i}(x)$

Caso 1 - soluzione 1

Idea: generalizzare le definizioni date nel caso delle partizioni

Partizione $\pi = \{A_1, \dots, A_n\}$	Covering $\gamma = \{B_1, \dots, B_n\}$
$\forall x \in X, \sum_{i=1}^n \chi_{A_i}(x) = 1$	$n(x) = \sum_{i=1}^n \chi_{B_i}(x)$ $\omega_{B_i}(x) = \frac{1}{n(x)} \chi_{B_i}(x)$

Caso 1 - soluzione 1

Idea: generalizzare le definizioni date nel caso delle partizioni

Partizione $\pi = \{A_1, \dots, A_n\}$	Covering $\gamma = \{B_1, \dots, B_n\}$
$\forall x \in X, \sum_{i=1}^n \chi_{A_i}(x) = 1$	$n(x) = \sum_{i=1}^n \chi_{B_i}(x)$ $\omega_{B_i}(x) = \frac{1}{n(x)} \chi_{B_i}(x)$ $\sum_{i=1}^n \omega_{B_i}(x) = 1$

Caso 1 - soluzione 1

Idea: generalizzare le definizioni date nel caso delle partizioni

Partizione $\pi = \{A_1, \dots, A_n\}$	Covering $\gamma = \{B_1, \dots, B_n\}$
$\forall x \in X, \sum_{i=1}^n \chi_{A_i}(x) = 1$	$n(x) = \sum_{i=1}^n \chi_{B_i}(x)$ $\omega_{B_i}(x) = \frac{1}{n(x)} \chi_{B_i}(x)$ $\sum_{i=1}^n \omega_{B_i}(x) = 1$
$ A_i = \sum_{x \in X} \chi_{A_i}(x)$	

Caso 1 - soluzione 1

Idea: generalizzare le definizioni date nel caso delle partizioni

Partizione $\pi = \{A_1, \dots, A_n\}$	Covering $\gamma = \{B_1, \dots, B_n\}$
$\forall x \in X, \sum_{i=1}^n \chi_{A_i}(x) = 1$	$n(x) = \sum_{i=1}^n \chi_{B_i}(x)$ $\omega_{B_i}(x) = \frac{1}{n(x)} \chi_{B_i}(x)$ $\sum_{i=1}^n \omega_{B_i}(x) = 1$
$ A_i = \sum_{x \in X} \chi_{A_i}(x)$	$m(B_i) = \sum_{x \in X} \omega_{B_i}(x)$

Caso 1 - soluzione 1

Idea: generalizzare le definizioni date nel caso delle partizioni

Partizione $\pi = \{A_1, \dots, A_n\}$	Covering $\gamma = \{B_1, \dots, B_n\}$
$\forall x \in X, \sum_{i=1}^n \chi_{A_i}(x) = 1$	$n(x) = \sum_{i=1}^n \chi_{B_i}(x)$ $\omega_{B_i}(x) = \frac{1}{n(x)} \chi_{B_i}(x)$ $\sum_{i=1}^n \omega_{B_i}(x) = 1$
$ A_i = \sum_{x \in X} \chi_{A_i}(x)$	$m(B_i) = \sum_{x \in X} \omega_{B_i}(x)$
$p(A_i) = \frac{ A_i }{ X }$	

Caso 1 - soluzione 1

Idea: generalizzare le definizioni date nel caso delle partizioni

Partizione $\pi = \{A_1, \dots, A_n\}$	Covering $\gamma = \{B_1, \dots, B_n\}$
$\forall x \in X, \sum_{i=1}^n \chi_{A_i}(x) = 1$	$n(x) = \sum_{i=1}^n \chi_{B_i}(x)$ $\omega_{B_i}(x) = \frac{1}{n(x)} \chi_{B_i}(x)$ $\sum_{i=1}^n \omega_{B_i}(x) = 1$
$ A_i = \sum_{x \in X} \chi_{A_i}(x)$	$m(B_i) = \sum_{x \in X} \omega_{B_i}(x)$
$p(A_i) = \frac{ A_i }{ X }$	$p(B_i) = \frac{m(B_i)}{ X }$

Entropia di un covering

$$H(\gamma) = - \sum_{i=1}^N p(B_i) \log p(B_i) = \log |X| - \frac{1}{|X|} \sum_{i=1}^N m(B_i) \log m(B_i)$$

Co-entropia

$$E(\gamma) = \frac{1}{|X|} \sum_{i=1}^N m(B_i) \log m(B_i)$$

Entropia di un covering

$$H(\gamma) = - \sum_{i=1}^N p(B_i) \log p(B_i) = \log |X| - \frac{1}{|X|} \sum_{i=1}^N m(B_i) \log m(B_i)$$

Co-entropia

$$E(\gamma) = \frac{1}{|X|} \sum_{i=1}^N m(B_i) \log m(B_i)$$

Problema: monotonia entropia-ordine non preservata!!

Entropia di un covering

$$H(\gamma) = - \sum_{i=1}^N p(B_i) \log p(B_i) = \log |X| - \frac{1}{|X|} \sum_{i=1}^N m(B_i) \log m(B_i)$$

Co-entropia

$$E(\gamma) = \frac{1}{|X|} \sum_{i=1}^N m(B_i) \log m(B_i)$$

Problema: monotonia entropia-ordine non preservata!!

Osservazione: non è l'entropia di un fuzzy set

(ad esempio De Luca, Termini, "A definition of a Nonprobabilistic Entropy in the Setting of Fuzzy Sets Theory", *Information and control*, 20, 301-312, 1972)

Caso 1 - soluzione 2 di Liang e Xu

Covering $\gamma = \{B_1, \dots, B_n\}$

Caso 1 - soluzione 2 di Liang e Xu

Covering $\gamma = \{B_1, \dots, B_n\}$

$$p_{LX}(B_i) := \frac{|B_i|}{|X|}$$

Caso 1 - soluzione 2 di Liang e Xu

Covering $\gamma = \{B_1, \dots, B_n\}$

$$p_{LX}(B_i) := \frac{|B_i|}{|X|}$$

NB: In generale non vale $\sum_{i=1}^N p_{LX}(B_i) = 1$
(Si possono modificare le cose in modo che valga... soluzione 3!)

Caso 1 - soluzione 2 di Liang e Xu

Covering $\gamma = \{B_1, \dots, B_n\}$

$$p_{LX}(B_i) := \frac{|B_i|}{|X|}$$

NB: In generale non vale $\sum_{i=1}^N p_{LX}(B_i) = 1$
(Si possono modificare le cose in modo che valga... soluzione 3!)

Pseudo entropia

$$H_{LX}(\gamma) = - \sum_{i=1}^N p_{LX}(B_i) \log p_{LX}(B_i) = \left(\sum_{i=1}^N |B_i| \frac{\log |X|}{|X|} \right) - \frac{1}{|X|} \sum_{i=1}^N |B_i| \log |B_i|$$

Caso 1 - soluzione 2 di Liang e Xu

Covering $\gamma = \{B_1, \dots, B_n\}$

$$p_{LX}(B_i) := \frac{|B_i|}{|X|}$$

NB: In generale non vale $\sum_{i=1}^N p_{LX}(B_i) = 1$
(Si possono modificare le cose in modo che valga... soluzione 3!)

Pseudo entropia

$$H_{LX}(\gamma) = - \sum_{i=1}^N p_{LX}(B_i) \log p_{LX}(B_i) = \left(\sum_{i=1}^N |B_i| \frac{\log |X|}{|X|} \right) - \frac{1}{|X|} \sum_{i=1}^N |B_i| \log |B_i|$$

Pseudo co-entropia (è la def. di entropia di L.X.)

$$E_{LX}(\gamma) := \frac{1}{|X|} \sum_{i=1}^N |B_i| \log |B_i|$$

Covering e ordine

S relazione di similarità

Ordine tra covering

Covering e ordine

S relazione di similarità

Ordine tra covering

$$\gamma \leq_s \delta \quad \text{iff} \quad \forall x : \mathcal{S}_\gamma(x) \subseteq \mathcal{S}_\delta(x)$$

Covering e ordine

S relazione di similarità

Ordine tra covering

$$\gamma \leq_s \delta \quad \text{iff} \quad \forall x : S_\gamma(x) \subseteq S_\delta(x)$$

Ordine stretto

$$\gamma <_s \delta \quad \text{iff} \quad \forall x : S_\gamma(x) \subseteq S_\delta \quad \text{and} \quad \exists x \in X \quad \text{such that} \quad S_\gamma(x) \neq S_\delta(x)$$

Covering e ordine

S relazione di similarità

Ordine tra covering

$$\gamma \leq_s \delta \quad \text{iff} \quad \forall x : \mathcal{S}_\gamma(x) \subseteq \mathcal{S}_\delta(x)$$

Ordine stretto

$$\gamma <_s \delta \quad \text{iff} \quad \forall x : \mathcal{S}_\gamma(x) \subseteq \mathcal{S}_\delta \quad \text{and} \quad \exists x \in X \quad \text{such that} \quad \mathcal{S}_\gamma(x) \neq \mathcal{S}_\delta(x)$$

Theorem (LX00)

$$\gamma \leq_s \delta \text{ implica } E_{LX}(\gamma) \leq E_{LX}(\delta)$$

Covering e ordine

S relazione di similarità

Ordine tra covering

$$\gamma \leq_s \delta \quad \text{iff} \quad \forall x : S_\gamma(x) \subseteq S_\delta(x)$$

Ordine stretto

$$\gamma <_s \delta \quad \text{iff} \quad \forall x : S_\gamma(x) \subseteq S_\delta \quad \text{and} \quad \exists x \in X \quad \text{such that} \quad S_\gamma(x) \neq S_\delta(x)$$

Theorem (LX00)

$$\gamma \leq_s \delta \text{ implica } E_{LX}(\gamma) \leq E_{LX}(\delta)$$

$$\gamma <_s \delta \text{ implica } E_{LX}(\gamma) < E_{LX}(\delta)$$

Covering e ordine

S relazione di similarità

Ordine tra covering

$$\gamma \leq_s \delta \quad \text{iff} \quad \forall x : S_\gamma(x) \subseteq S_\delta(x)$$

Ordine stretto

$$\gamma <_s \delta \quad \text{iff} \quad \forall x : S_\gamma(x) \subseteq S_\delta \quad \text{and} \quad \exists x \in X \quad \text{such that} \quad S_\gamma(x) \neq S_\delta(x)$$

Theorem (LX00)

$$\gamma \leq_s \delta \text{ implica } E_{LX}(\gamma) \leq E_{LX}(\delta)$$

$$\gamma <_s \delta \text{ implica } E_{LX}(\gamma) < E_{LX}(\delta)$$

$$A \subset B \subseteq Att \text{ implica } E_{LX}(\gamma(A)) \leq E_{LX}(\gamma(B))$$

Definition

Let $\mathcal{K}^{(t_0)}(X)$ and $\mathcal{K}^{(t_1)}(X)$ be two incomplete information systems based on the same triple $\langle X, Att, val \rangle$ and characterized by two different information maps

$$F^{(t_0)} : X \times Att \mapsto val \quad \text{and} \quad F^{(t_1)} : X \times Att \mapsto val$$

Definition

Let $\mathcal{K}^{(t_0)}(X)$ and $\mathcal{K}^{(t_1)}(X)$ be two incomplete information systems based on the same triple $\langle X, Att, val \rangle$ and characterized by two different information maps

$$F^{(t_0)} : X \times Att \mapsto val \quad \text{and} \quad F^{(t_1)} : X \times Att \mapsto val$$

There is a **monotonic increase of information** iff

$$t_0 \leq t_1 \quad \text{and} \quad \forall (x, a) \ (F^{(t_0)}(x, a) \neq * \Rightarrow F^{(t_1)}(x, a) = F^{(t_0)}(x, a))$$

Definition

Let $\mathcal{K}^{(t_0)}(X)$ and $\mathcal{K}^{(t_1)}(X)$ be two incomplete information systems based on the same triple $\langle X, Att, val \rangle$ and characterized by two different information maps

$$F^{(t_0)} : X \times Att \mapsto val \quad \text{and} \quad F^{(t_1)} : X \times Att \mapsto val$$

There is a **monotonic increase of information** iff

$$t_0 \leq t_1 \quad \text{and} \quad \forall (x, a) \ (F^{(t_0)}(x, a) \neq * \Rightarrow F^{(t_1)}(x, a) = F^{(t_0)}(x, a))$$

We write **$\mathcal{K}^{(t_0)}(X) \preceq \mathcal{K}^{(t_1)}(X)$**

In breve: sostituisco $*$ con valori definiti

In breve: sostituisco $*$ con valori definiti

Theorem

$\mathcal{K}^{(t_0)}(X) \preceq \mathcal{K}^{(t_1)}(X)$ *implica* $\gamma(t_1) \leq_s \gamma(t_0)$

Quindi

$$E_{LX}(\gamma(t_1)) \leq E_{LX}(\gamma(t_0))$$

The end