

Teoria dell'informazione

Perchè

- Operatori umani come macchine che codificano e trasmettono informazione
- La teoria dell'informazione rende disponibile una nozione precisa di trasmissione dell'informazione
- Rende possibile determinare la difficoltà di un dato compito (**incertezza**)
- Rende possibile determinare l'efficienza dell'elaborazione (**informazione trasmessa**)

Informazione come incertezza - 1

- Caso 1: presentazione ripetuta dello stesso stimolo. Incertezza nulla
- Caso 2: due stimoli s1 ed s2 possono essere presentati, con la stessa probabilità.
Se vedo s1, di quanto si riduce l'incertezza=quanta informazione rende disponibile l'occorrenza di s1?
- Caso 4: ci sono quattro stimoli, s1, s2, s3, s4, equiprobabili
Se vedo s2, di quanto si riduce l'incertezza=quanta informazione rende disponibile l'occorrenza di s2?

Informazione come incertezza - 2

- Idealmente (e con probabilità uguali) più sono le possibilità (stimoli) più l'apparire di uno di essi riduce l'incertezza iniziale=più l'apparire di uno di essi fornisce informazione.
- Probabilità non uguali:
 'oggi il sole tramonerà'
 'domani l'Inghilterra dichiarerà guerra agli USA'
- Quali dei due è più informativo (riduce maggiormente l'incertezza)?

Informazione come incertezza - 3

- Eventi con probabilità più bassa sono più informativi = la loro occorrenza diminuisce l'incertezza in modo maggiore
- La misura dell'informazione (incertezza) deve considerare:
 - Il numero di eventi in gioco
 - Le loro probabilità
 - Il contesto in cui gli eventi occorrono

Misura dell'informazione - 1

- Paradigma domanda-risposta
- La riduzione di incertezza (=informazione) è pari al numero di domande sì/no che si dovrebbero porre per ridurre l'incertezza.
- Elezioni: ci sono due candidati, X ed Y
- Domande necessarie:
 - Ha vinto X?
 - Informazione dell'evento (riduzione dell'incertezza)) = **1 bit**

Misura dell'informazione - 2

- Elezioni: quattro candidati: FI, AN, DS, Margherita
- Domande
 - Ha vinto il Polo?
 - Ha vinto il candidato più di centro?
- Informazione=riduzione dell'incertezza= 2 bits

- La misura dell'informazione deve essere direttamente proporzionale al numero di possibilità

Misura dell'informazione - 3

- Eventi poco probabili sono altamente informativi (forte riduzione dell'incertezza)
- Eventi molto probabili sono poco informativi (bassa riduzione dell'incertezza)

- La quantità d'informazione è inversamente proporzionale alla probabilità dell'evento

Misura dell'informazione - 4

- Misura dell'informazione=incertezza=Entropia

$$H = -\sum_{i=1}^n p_i \log_2 p_i$$

Esempio - 1

	A	B	C	D
p_i	0.5	0.25	0.125	0.125
$\log_2 p_i$	-1	-2	-3	-3

H=1.75 bits

Esempio - 2

	A	B	C	D
P_i	0.25	0.25	0.25	0.25
$\log_2 p_i$	-2	-2	-2	-2

$H=2$ bits

Quindi

- Dati n eventi possibili, la massima incertezza/ informazione/ entropia si ha quando sono equiprobabili
- Se alcuni hanno probabilità maggiori di altri, l'incertezza globale/ informazione/ entropia diminuisce

Serie di eventi

- La lettera *u*.
- La sequenza ABABABAB
- Cosa viene dopo?

- Probabilità condizionali

Ridondanza - 1

- Il numero massimo di eventi fissa un limite superiore all'incertezza/ quantità di informazione (entropia).
- Probabilità diseguali o restrizioni contestuali / sequenziali, riducono l'incertezza/ informazione / entropia.
- Ridondanza: misura della riduzione di informazione/ incertezza/ entropia dovuta al contesto/ probabilità disuguali

Ridondanza - 2

- Esempio: il linguaggio umano
 - Le lettere non sono equiprobabili
 - Restrizioni sulle occorrenze (q è sempre seguita da u)
 - La sintassi di una lingua determina
- Contributo della sintassi

Ridondanza - 3

- Ridondanza percentuale di un insieme di stimoli

$$(1 - H_e / H_{\max}) * 100$$

- La ridondanza dell'alfabeto inglese è 68%

Trasmissione dell'informazione

- Informazione della sorgente
- Informazione ricevuta
- Informazione trasmessa
- Informazione perduta
- Rumore

- Il caso del dattilografo.

Esempio

	A	B	C	D	
A	2 (0,25)				2 (0,25)
B		2 (0,25)			2 (0,25)
C			2 (0,25)		2 (0,25)
D				2 (0,25)	2 (0,25)
	2 (0,25)	2 (0,25)	2 (0,25)	2 (0,25)	

	A	B	C	D	
A	1 (.125)			1 (.125)	2 (0,25)
B		1 (.125)	1 (.125)		2 (0,25)
C	1 (.125)			1 (.125)	2 (0,25)
D		1 (.125)	1 (.125)		2 (0,25)
	2 (0,25)	2 (0,25)	2 (0,25)	2 (0,25)	

$$H_T = H_S + H_R - H_{SR} = 2 + 2 - 2 = 2 \text{ bits}$$

$$H_T = H_S + H_R - H_{SR} = 2 + 2 - 3 = 1 \text{ bit}$$

Trasmissione dell'informazione

- $T = H_s + H_r - H_{sr}$
- Quantita' media di informazione (per stimolo) che passa nella risposta
- Entropia media della (sequenza di) risposta che e' spiegata dall'entropia degli stimoli
- Riconoscimento perfetto: $H_s = H_r = H_{sr}$
- Entropia massima: stimoli e risposte indipendenti
 $H = H_s + H_r$

Trasmissione dell'informazione

- $L = H_{sr} - H_r$
- $L =$ informazione perduta=entropia (media) degli stimoli non spiegata dalle risposte
- $L + T = H_s$
- $G = H_{sr} - H_s$
- $G =$ entropia delle risposte non spiegata dallo stimolo=contributo del soggetto
- $G + T = H_r$