

Intelligenza artificiale, informatica, linguistica, scienze umane e molte altre discipline concorrono a formare il campo che oggi si definisce “**informatica umanistica**”

Informatica

- **Informazione+automatica** = informatica
 - “Trattamento automatico dell’informazione”
 - La nozione di informazione è **ambigua** e usata in modi molto diversi nei diversi ambiti
- Che cos’è l’**informazione**?
 - L’informazione è un contenuto
 - L’informazione non è un contenuto, è la semplice presenza di una scelta
- Che cosa significa “**automatico**”?
 - Non richiede l’intervento umano (nel dettaglio)
 - Meccanico (deterministico)

La teoria dell'informazione

- Claude Shannon e Warren Weaver, 1949, *La teoria matematica della comunicazione*
- Dalla crittografia ai codici
- Problemi di Shannon:
 - trasmissione materiale di informazioni su un canale
 - misurare la capacità di un canale
 - stabilire quale sia il miglior sistema di codifica dei messaggi per garantire la ricezione



Claude
Shannon
(1916-2001)

L'informazione come incertezza

- La **trasmissione** è qualcosa di materiale che consiste nel fare in modo che una qualunque variazione di uno stato fisico (onde sonore, porzioni grafiche, o tattili, ma anche in linea di principio olfattive o gustative) passi su un canale e venga ricevuta correttamente da un apparecchio di ricezione di tali variazioni
- L'informazione è **indipendente dal contenuto** (asemantica)
- L'informazione corrisponde a una **scelta** tra due possibilità equiprobabili
- L'informazione è connessa con l'incertezza, con la sorpresa.

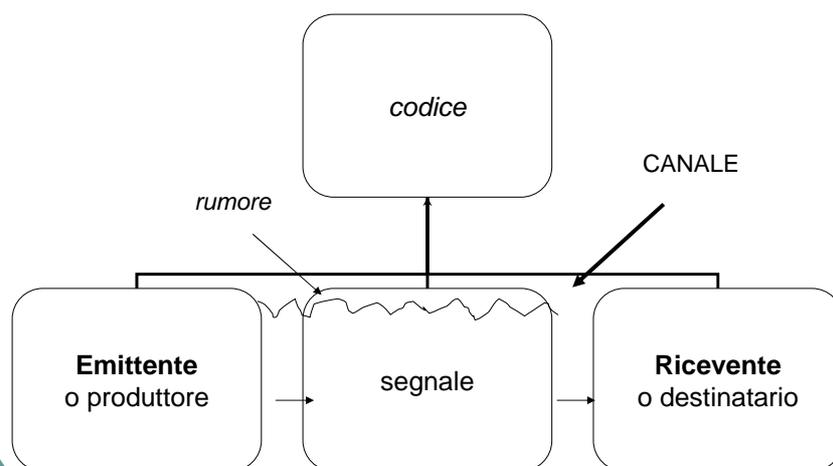
L'informazione come entropia

- L'informazione consiste nella rimozione dell'incertezza della situazione iniziale
- **Maggiore è l'incertezza, maggiore sarà l'informazione portata dal messaggio e minore sarà lo stato di incertezza del ricevente al termine del messaggio**
- L'informazione non ha un valore assoluto ma solamente **relazionale**
- La nozione matematica (e teorica) di informazione in questo senso coincide con la nozione fisica di **ENTROPIA**

$$H(X) \equiv - \sum_x P(x) \log_2 [P(x)]$$

- **grandezza termodinamica il cui aumento misura la diminuzione dell'energia posseduta da un sistema isolato**
- **Tendenza al disordine, caos**

Lo schema della comunicazione



Gli elementi della comunicazione

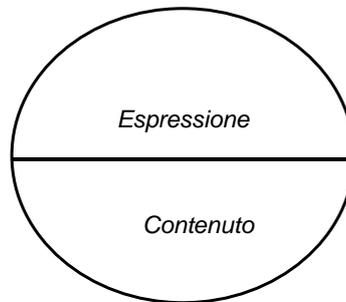
- la **sorgente** o **produttore** o **emittente**, ossia il dispositivo che produce una variazione di stato fisico;
- il **ricevente** o **destinatario**, ossia il dispositivo che rileva le variazioni di stato fisico;
- il **messaggio**, ossia l'insieme delle variazioni di stato fisico emesse dal produttore (collegabili a una serie di contenuti anche complessi che però non entrano nel processo di trasmissione);
- il **canale**, ossia il mezzo fisico su cui viene trasmesso;
- la **codifica**, ossia il processo attraverso il quale sono formati i messaggi e la **decodifica** ossia il processo di ricostruzione del messaggio da parte del ricevente;
- il **rumore**, ossia il possibile disturbo che può intervenire sul canale

Il codice

- È formato da:
- Un **ALFABETO** (composto dai simboli elementari del linguaggio)
 - Gli elementi dell'alfabeto sono in *opposizione reciproca*
- Una **GRAMMATICA** (una serie di regole di combinazione dei simboli dell'alfabeto)
- *Eventualmente* da una serie di associazioni di ogni combinazione di simboli a un significato

Codici e linguaggi

- Il codice è simile a un linguaggio, ma non lo è
 - Un codice ha un piano costituito da una sua struttura
- Il linguaggio è invece composto dall'associazione di due piani (è *biplanare*) detti:
 - Piano dell'espressione
 - Piano del contenuto
- La Teoria dell'informazione NON tratta di un linguaggio, tratta di strutture monoplanari



SEGNO

Teorema fondamentale per un canale discreto con disturbo di Shannon

- è sempre possibile una codifica che permetta la trasmissione di un messaggio senza il prezzo di una velocità troppo bassa
- Nelle comunicazioni, dice Shannon,
«vi sono, comunque, dei modi di trasmettere le informazioni i quali sono ottimali nel combattere il disturbo»
- I diversi mezzi per combattere il rumore sul canale sono chiamati **RIDONDANZA**
- La *ridondanza* infatti è una codificazione del messaggio che ripetendo alcune porzioni oppure variando le probabilità di occorrenza delle unità permette di garantire la corretta ricezione del messaggio

Rappresentazione analogica e digitale

| Analogico | Digitale |
|--|--|
| Rappresentazione per analogia con altre grandezze fisiche | Può adottare una numerazione anche binaria per la rappresentazione |
| Difficile rimuovere gli effetti di rumore | Semplice rimuovere il rumore della rappresentazione |
| Nel caso di uso discretizzato perde la sua maggiore aderenza al fenomeno rappresentato | È pensata per essere usata in un contesto discreto |
| Vinile, orologi a lancette ecc. | CD, orologi digitali, ecc. |

La rappresentazione digitale dell'informazione

- La codifica è una rappresentazione dell'informazione
- Un codice estremamente preciso è quello che usa i **numeri naturali**
- Una codifica si dice **binaria** quando si usano solo due cifre, 0 e 1
- Digit (inglese) = cifra; numerique (francese)= digitale
- I codici BINARI sono digitali

L'unità di misura dell'informazione digitale: il bit

- Il **bit** (binary digit): unità di misura dell'informazione digitale
- 1 **bit**: la quantità di informazione fornita dalla scelta tra due diverse alternative equiprobabili
- 1 **bit**: unità minima dell'informazione digitale, è una singola cifra, 0 o 1
- Ogni bit può avere due possibilità: 0 o 1
- 1 byte = 8 bit
- 1 byte può codificare 2^8 diverse possibilità: 256 combinazioni diverse di 0 e 1

Il digitale

Qualunque tipo di messaggio può essere trasmesso attraverso un codice binario e dunque digitale

- Immagini, suoni, testi possono essere
 - **Digitalizzati** attraverso diversi processi (scannerizzazione, campionatura, riconoscimento ottico, ecc.)
 - **Trasmessi**
 - **Conservati** su un unico supporto
- Il costo di tali operazioni si misura sempre in *bit*

Perché è importante la teoria dell'informazione?

- È il fondamento della nozione di *digitale*
- è alla base di tutte le applicazioni informatiche, ivi compresa la trasmissione via Internet
- centralità della dimensione statistico-probabilistica rispetto alla dimensione logico-formale

